CHAPTER 4

UNIT IV

Security Requirements

Syllabus :

At the end of this unit, you should be able to understand and comprehend the following syllabus topics :

- IP Security
  - Introduction
  - Architecture
  - IPV6
  - IPv4
  - IPSec protocols and Operations
    - AH Protocol
    - ESP Protocol
    - ISAKMP Protocol
    - Oakleykey determination Protocol
  - VPN
- Web Security
  - Introduction
  - Secure Socket Layer (SSL)
    - SSL Session and Connection
    - SSL Record Protocol
    - Change Cipher Spec Protocol
    - Alert Protocol
    - Handshake Protocol
  - Secure Electronic Transaction (SET)
- Electronic Mail Security
  - Introduction
  - Pretty Good Privacy (PGP)
  - MIME
  - S/MIME
  - Comparison

4.1 IP Security

Q. Differentiate between IPV4 and IPV6. SPPU - May 19

(May 19, 4 Marks)

– IP stands for Internet Protocol. IP defines a set of protocols that can be used for communication between any two devices on the network. A network protocol is a standard set of rules that determines how systems will communicate across networks.

- Two different systems that use the same protocol can communicate and understand each other very similar to how two people can communicate and understand each other by using the same language.
- IP provides addressing and routing mechanisms for each packet of data that needs to move across the network. Each device on the network must have a unique IP address to communicate with any other device on the network.

Note : It is assumed that you have a general understanding of computer networking. While this section does not dive deeper into computer networks, it focuses on specific security topics around networking.

4.1.1 IPv4

- IPv4 is IP version 4. This is the most common IP addressing scheme used today despite certain challenges. It is 32-bit long and thus has an address space of 2^{32} = 4,294,967,296. This means you can maximally have 4,294,967,296 (approximately 4.3 billion) IPv4 addresses. There are many more devices than the number 4,294,967,296. Fig. 4.1.1 shows an outline of the IPv4 header.

Offsets Octet	Octet Bit	0								1								2								3							
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
0	0	Version				IHL				DSCP						ECN		Total Length															
4	32	Identification																Flags		Fragment Offset													
8	64	Time To Live								Protocol								Header Checksum															
12	96	Source IP Address																															
16	128	Destination IP Address																															
20	160	Options (if IHL > 5)																															
24	192																																
28	224																																
32	256																																

Fig. 4.1.1 : IPv4 header

- An example of IPv4 address looks like 121.56.78.214.

4.1.2 IPv6

- IPv6 was created to address the limitation of IPv4 to have only 4,294,967,296 IP addresses due to 32-bit length. IPv6 more or less provides the similar addressing and routing capabilities but one core difference between IPv4 and IPv6 is the address space. IPv6 address is 128-bit long and thus you can have 2^{128} IPv6 addresses!
- Fig. 4.1.2 shows an outline of the IPv6 header.

Offsets Octet	Octet Bit	0								1								2								3							
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
0	0	Version				Traffic Class												Flow Label															
4	32	Payload Length																Next Header								Hop Limit							
8	64	Source IP Address																															
12	96																																
16	128																																
20	160																																
24	192	Destination IP Address																															
28	224																																
32	256																																
36	288																																

Fig. 4.1.2 : IPv6 header

- An example of IPv6 address looks like 2001:0:9d38:6abd:2c37:10da:8554:4234.

4.1.3 Internet Protocol Security (IPSec)

SPPU - May 19

Q. Discuss working of IPSec. What are the benefits of IPSec. (May 19, 6 Marks)

Definition :IPSec is a suite of protocols that protects IP traffic.

- IP does not have any integrated security mechanisms by itself and hence IPSec (short form for IP Security) is additionally used to provide security for IP traffic.
- The IPSec suite consists of following security protocols.

Table 4.1.1 : Security protocols provided by IPSec

Sr. No.	Protocol Name	Functionality Provided
1.	Authentication Header (AH)	Data Integrity Data Origin Authentication Protection from replay attacks
2.	Encapsulating Security Payload (ESP)	Confidentiality Data Origin Authentication Data Integrity
3.	Internet Security Association and Key Management Protocol (ISAKMP)	Framework for Authentication and Key Exchange
4.	Internet Key Exchange (IKE)	Authenticated keying material for use with ISAKMP

- Here IPSec is a framework. It does not mandate which hashing and encryption algorithms should be used or how keys should be exchanged between the communicating devices. Key management can be handled manually or automated by a key management protocol such as ISAKMP.

1. Security Association

- Security Association (SA) is a fundamental concept with respect to IPSec.

Definition : Security Association holds several information that determines how security services would be consumed by the communicating devices.

- IPSec provides many options for performing security services such as encryption, integrity and authentication. The network devices that wish to establish an IPSec connection, must negotiate and arrive at exactly which algorithms and parameters to use for the chosen IPSec security services. The security association is a mechanism to hold all the agreed terms (algorithms, parameters, etc.) for a given IPSec communication session.

2. Modes of operation

IPSec can work in two modes.

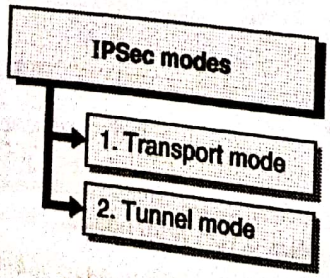


Fig. 4.1.3 : IPSec Modes

A. Transport Mode

In this mode, only the payload (data) part of the information is protected. The addressing and routing information is not protected. It is like a sealed envelope with address on it. The message inside the envelop is protected whereas the source and destination addresses are not.

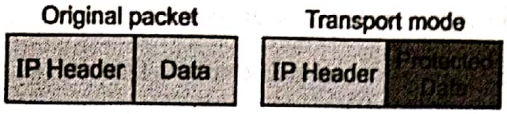


Fig. 4.1.4 : Transport Mode

B. Tunnel Mode

In this mode, both the payload (data) as well as the addressing information is protected. In this mode, the entire packet is protected, and a new IP header is added by IPSec. The original IP header information along with the payload information is protected. Tunnel mode provides more security than the transport mode.

Fig. 4.1.5 : Tunnel Mode

3. Applications / Benefits / Usage of IPSec

I. Establish Virtual Private Network (VPN)

IPSec is predominantly used to establish VPN connection. VPN connections are generally used to access private networks over the internet. For example, you can access your college or your organization's network from home over the internet.

II. Connecting two or more branch networks

IPSec can be used to extend or connect branch networks. For example, if you have two branches of office each using its own network, the branches can be connected using IPSec. The network traffic then can securely move between the branches.

III. General security benefits

IPSec adds general security benefits to the core IP protocol. It provides benefits such as data confidentiality, data integrity, data origin authentication and protection from several attacks on the core IP protocol.

4. How does IPSec work?

Overall, communication over IPSec has 5 broad steps.

1. **Initiate IPSec process :** IPSec communication begins with the identification of traffic that requires IPSec security.
2. **IKE Phase 1 :** In this phase, the IKE SAs are negotiated and agreed.
3. **IKE Phase 2 :** In this phase, next set of SAs for actual data transfer are negotiated and agreed.
4. **Data Transfer :** Data is transferred between the communicating entities.
5. **Termination :** The IPSec connection is terminated once the data transfer is complete.

4.1.4 Authentication Header (AH)

✍ ***Definition :*** *The Authentication Header (AH) protocol provides data integrity and data original (source address) authentication over the network communication.*

- It also provides replay protection. It does not provide encryption.
- AH calculates the Integrity Check Value (ICV) over non-changing fields of the IP header:
  - Next Header
  - Payload Len
  - Reserved
  - Security Parameter Index (SPI)
  - Sequence number
  - Padding bytes
- ICV is a hash value which is often computed using SHA-1 or other hashing algorithms.
- In the transport mode, Fig. 4.1.6 shows the simplistic diagram of before and after applying AH.

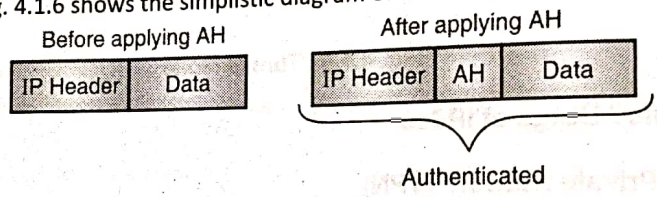


Fig. 4.1.6

- In the tunnel mode, Fig. 4.1.7 shows the simplistic diagram of before and after applying AH.

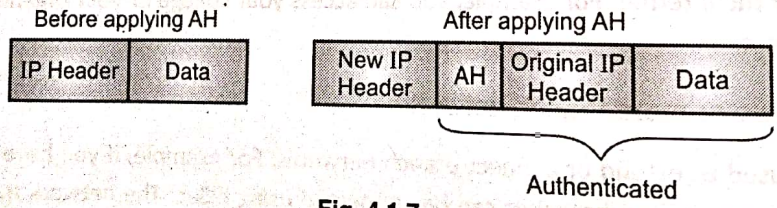


Fig. 4.1.7

- As we discussed earlier on hash values, hash values provide integrity. AH uses hashing algorithms to find out hash value of the IP header and attaches it with the IP header. This way it not only provides data integrity (since hash is calculated on payload as well) but also data origin integrity or authentication (since source address is part of the IP header as well).

4.1.5 Encapsulating Security Payload (ESP)

✍ ***Definition :*** *The ESP protocol is designed to provide confidentiality (through encryption), data integrity and data origin authentication over network communication.*

- ESP can be applied with AH or without AH. Here ESP can itself provide integrity. It does not need AH for integrity. You have an option to additionally calculate integrity using AH. ESP in transport mode encrypts the actual payload (data) so that it cannot be read by an unauthorized entity. In tunnel mode, the IP header information is encrypted as well.
- If you choose integrity service, the Integrity Check Value (ICV) is calculated on the following fields in the IP header.
  - Security Parameter Index (SPI)

- Sequence Number
- Payload Data
- ESP trailer

If you choose confidentiality service, the ciphertext consists of the following fields in the IP header.

- Payload (Data)
- ESP trailer

In the transport mode, Fig. 4.1.8 shows the simplistic diagram of before and after applying ESP.

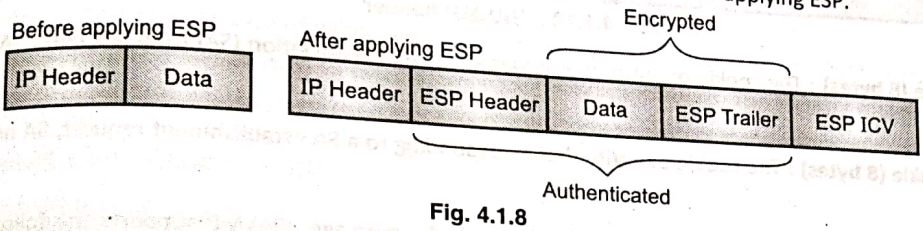


Fig. 4.1.8

In the tunnel mode, Fig. 4.1.9 shows the simplistic diagram of before and after applying ESP.

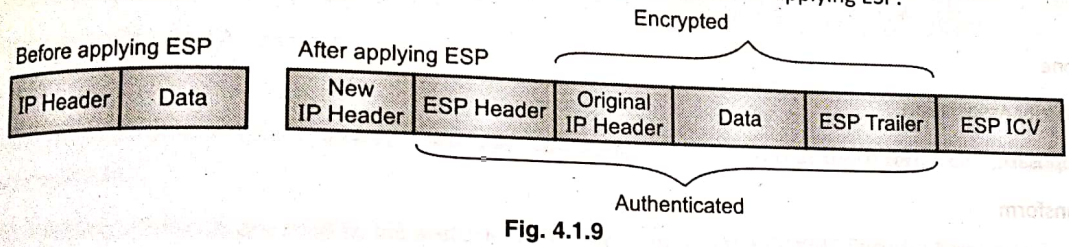


Fig. 4.1.9

4.1.6 Internet Security Association and Key Management Protocol (ISAKMP)

SPPU – May 19

Q. Explain ISAKMP protocol of IPSec with header format. (May 19, 7 Marks)

Definition : ISAKMP provides a framework for authentication and key exchange.

ISAKMP does not define the exact algorithms to be used. It is just a framework within which various exchange protocols can work.

ISAKMP defines the procedures for

- Authenticating communication devices
- Creation and management of Security Associations (SA)
- Key generation techniques
- Threat mitigation

ISAKMP is distinct from key exchange protocols in order to cleanly separate the details of security association management (and key management) from the details of key exchange. There may be many different key exchange protocols, each with different security properties. However, a common framework is required for agreeing to the format of SA attributes, and for negotiating, modifying, and deleting SAs. ISAKMP serves as this common framework.

Fig. 4.1.10 shows a simplistic diagram of ISAKMP header.

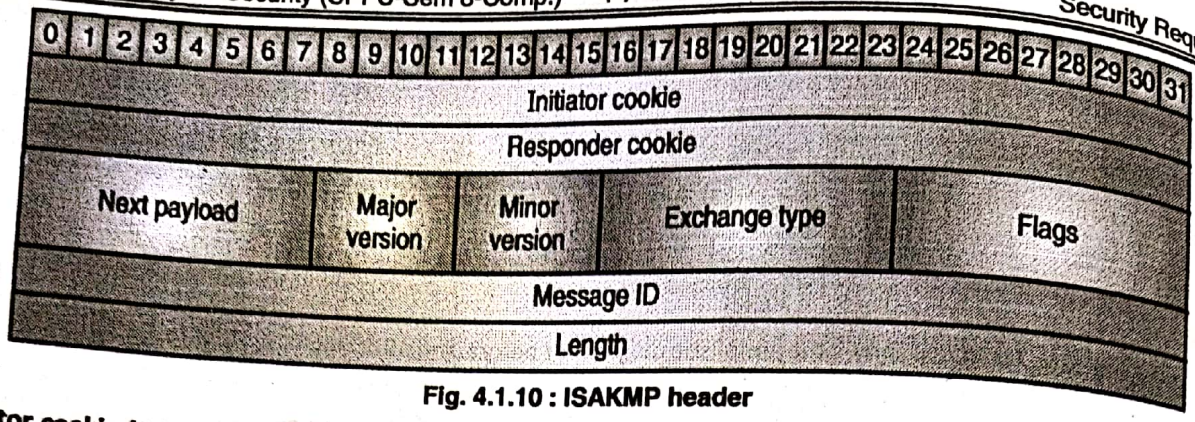


Fig. 4.1.10 : ISAKMP header

1. **Initiator cookie (8 bytes) :** The cookie of entity that initiated Security Association (SA) establishment, SA notification, or SA deletion.
2. **Responder cookie (8 bytes) :** The cookie of entity that is responding to a SA establishment request, SA notification, or SA deletion.
3. **Next Payload (1 byte) :** Indicates the type of first payload in the message. ISAKMP supports the following payload types:
   - None
   - Security Association
   - Proposal
   - Transform
   - KeyExchange
   - Identification
   - Certificate
   - Certificate Request
   - Hash
   - Signature
   - Nonce
   - Notification
   - Delete
   - VendorID
   - NAT Discovery Payload
   - NAT Original Address Payload
   - Reserved
   - PrivateUse
4. **Major version (4-bits) :** Major version of the ISAKMP protocol in use.
5. **Minor version (4-bits) :** Minor version of the ISAKMP protocol in use.

6. **Exchange Type (1 byte) :** The type of exchange in a given ISAKMP session. The primary difference between exchange types is the ordering of the messages and the payload ordering within each message.
7. **Message ID (4 bytes) :** The unique message identifier.
8. **Length (4 bytes) :** The length, in bytes, of the total message (header + payloads).

- ISAKMP offers two phases of negotiation.
  - **Phase 1 :** In the first phase, two entities agree on how to protect further negotiation traffic between themselves, establishing an ISAKMP SA.
  - **Phase 2 :** The second phase of negotiation is used to establish security associations for other security protocols. This second phase can be used to establish many security associations. The security associations established by ISAKMP during this phase can be used by a security protocol to protect many message/data exchanges.

4.1.7 Internet Key Exchange (IKE)

Definition : Internet Key Exchange (IKE) is the protocol used to set up a Security Association (SA) in the IPsec protocol suite.

- IPSec currently uses IKE version 2.
- Recall from our earlier discussion on security association – A security association is a set of negotiated terms (algorithms, parameters, etc.) between two communicating entities such that these terms can be used in successive communication.
- The following attributes are used by IKE and are negotiated as part of the ISAKMP Security Association:
  - Encryption algorithm
  - Hashing algorithm
  - Authentication method
  - Information about a group over which to do Diffie-Hellman exchange
- All of these attributes are mandatory and MUST be negotiated between the communicating entities. IKE supports the following attributes for negotiation.

Attribute For	Supported Attributes
Encryption algorithm	DES, IDEA, Blowfish, 3DES, CAST, RC5, AES
Hashing algorithm	MD5, SHA, TIGER
Authentication method	Pre-shared key, DSS Signature, RSA Signature, Encryption with RSA, Revised encryption with RSA
Group information	MODP (modular exponentiation group), ECP (elliptic curve group over GF[P]), EC2N (elliptic curve group over GF[2^N])

- IKE works in two phases.
- In Phase 1, following functions are carried out:

- Mutual authentication of the communicating entities.
- Negotiating cryptographic parameters.
- Creating session keys.

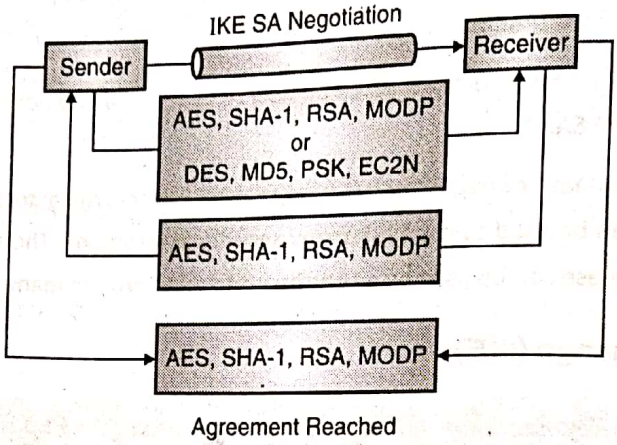


Fig. 4.1.11

– **In Phase 2,** an IPSec tunnel is negotiated by creating keying material for the IPSec tunnel to use (either by using the IKE phase one keys as a base or by performing a new key exchange).

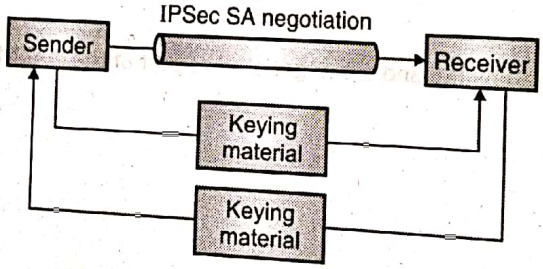


Fig. 4.1.12

4.1.8 OAKLEY Key Determination Protocol

Definition : *OAKLEY is a key determination protocol using which two authenticated parties can agree on secure and secret keying material.*

– It is based on the Diffie-Hellman key exchange algorithm. The OAKLEY protocol has compatibility with the ISAKMP protocol for managing security associations, user-defined abstract group structures for use with the Diffie-Hellman algorithm, key updates, and incorporation of keys distributed via out-of-band mechanisms.

– The OAKLEY protocol is used to establish a shared key with an assigned identifier and associated authenticated identities for the two parties. The name of the key can be used later to derive security associations for AH and ESP. At a high-level, there are three components of the key determination protocol.

1. Cookie exchange
2. Diffie-Hellman key exchange
3. Authentication

Note : Current IPsec implementations actually use IKEv2. But its predecessor, IKEv1, was based on the OAKLEY and Security Key Exchange Mechanism (SKEME) protocols.

4.2 VPN

SPPU – May 19

Q. What is VPN? (May 19, 3 Marks)

Organisations setup internal and private network for use by its authorised users. Its resources are not accessible from public network such as the Internet. Increasingly the task force is becoming global and remote. Physical presence to access the organisation resources is no more efficient. At the same time, the organisation cannot risk exposing its internal resources over the public network.

A Virtual Private Network (VPN) provides a solution for this scenario. It allows to establish a secure channel between the communicating parties over the public network, such as the Internet and facilitate secure connection between them.

A Virtual Private Network (VPN) is a data network that enables two or more parties to communicate securely across a public network by creating a private connection, or "tunnel," between them.

The authorised users can then securely access the private network over the public network. Physical presence within organisation premises is not required to access the private network. The entire traffic between the remote user and the private network is encrypted. IPSec can be one of the mechanisms for establishing a VPN connection.

4.2.1 Types of VPN

SPPU – May 19

Q. Explain types of VPN. (May 19, 3 Marks)

At a broad level, there are two types of VPN.

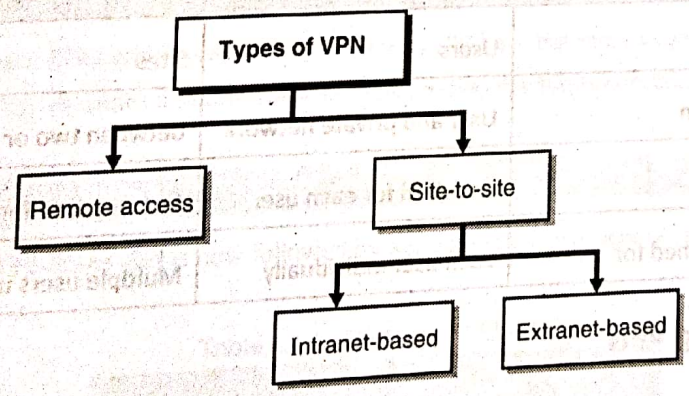


Fig. 4.2.1

1. **Remote Access VPN**

Remote Access VPN is setup for remote users. They can access the private network securely over the public network.

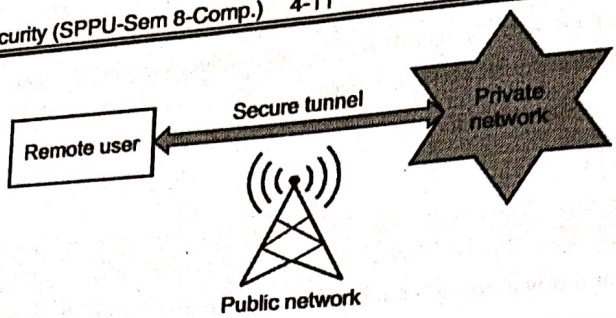


Fig. 4.2.2

2. **Site-to-Site VPN**

This is established by the organisation for connecting its multiple sites or branch offices so that the users can access the resources across the sites.

a. **Intranet-based site-to-site VPN** is used for organization's own sites.

b. **Extranet-based site-to-site VPN** is used for organization and its partners.

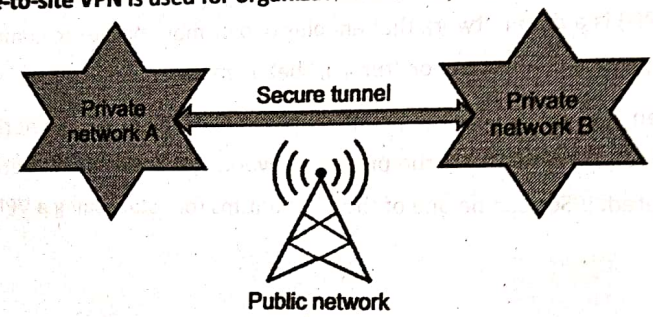


Fig. 4.2.3

Comparison between Remote Access and Site-to-Site VPN

Sr. No.	Comparison Attribute	Remote Access VPN	Site-to-Site VPN
1.	Setup For	Users	Sites
2.	Tunnel between	User and private network	Between two or more private networks
3.	VPN client	Required for each user	Not required for each user
4.	Tunnel established for	Each user individually	Multiple users use the same tunnel

4.2.2 Challenges of using VPN

VPN technologies have the following challenges.

1. **Interoperability :** There are various types of clients these days – Laptop, Desktop, Tablets, Mobile Phones, each running a variety of OS and applications. Ensuring that the VPN technology is compatible with all the possible client types if a major challenge.

2. **Installation and management of VPN clients and gateways :** The use of a remote access VPN requires that a VPN client be installed on each device. A site-to-site VPN requires installing and managing several VPN gateways. If any of

the VPN gateways is down, the users would not be able to access the remote resources.

3. **Client security :** VPN allows access to private network and remote resources. It is important to ensure that the clients accessing the private network are secure enough. If the client security is compromised, it might infect the private network as well.

4. **Requires strong authentication :** Since VPNs provide access to the private network, the user must be strongly authenticated to ensure that she is a legitimate user. You must use two-factor authentication to ensure that the user is authorized to use the private network.

4.3 Web Security

- World Wide Web or just web is a collection of web servers that run several websites that hold the desired information.
- The Internet as a whole is a collection of such servers and various communication devices and protocols. You mostly use browsers (such as Chrome, Firefox, Internet Explorer, Safari, etc.) or applications (for example, Mobile Apps) to browse the web and fetch the desired information or just complete a desired interaction such as making a purchase.
- Let me pause you here and ask a simple question. Don't you feel that your interaction with the Internet (which generally is an insecure and unsafe place) should be protected? For example, if you type your Facebook password, should it be available to everyone on the network?
- To make a purchase when you provide your bank account information, isn't that information very confidential and requires secure handling as you pass it through your device all the way to the website? Yes, I am sure you understand that your interaction with the web requires security. There is one protocol that we all need – SSL (obsolete now) followed by TLS (currently used). Let's learn about it.

4.4 Secure Socket Layer (SSL)

- The Secure Sockets Layer (SSL) and Transport Layer Security (TLS) are the most widely use web security protocol. It is essentially a protocol that provides a secure channel between two machines operating over the Internet or an internal network.

Definition - SSL is a cryptographic protocol designed to protect communication between two entities.

- SSL underwent several revisions and is now followed by a more secure protocol called TLS. Table 4.4.1 shows a quick version history of SSL/TLS.

Table 4.4.1 : History of SSL/TLS

Protocol	Published	Status
SSL 1.0	Unpublished	Unpublished
SSL 2.0	1995	Obsoleted in 2011
SSL 3.0	1996	Obsoleted in 2015
TLS 1.0	1999	To be obsolete in 2020

Protocol	Published	Status
TLS 1.1	2006	To be obsolete in 2020
TLS 1.2	2008	Currently good
TLS 1.3	2018	Currently good

Goals of SSL

1. **Cryptographic Security :** Establish and provide a secure connection between two parties.
2. **Interoperability :** Two unrelated applications should be able to establish SSL connection.
3. **Extensibility :** Provides a framework for using various algorithms and methods without changing the protocol.
4. **Efficiency :** Performance enhancement mechanism to avoid overloading the system when protocol is in use.

4.4.1 Overview of SSL Protocol

- SSL protocol works in layers. At each layer, messages may include fields for length, description, and content. SSL takes messages to be transmitted, fragments the data into manageable blocks, optionally compresses the data, applies a MAC, encrypts, and transmits the result. On the other side, received data is decrypted, verified, decompressed, and reassembled and then delivered to higher level clients.

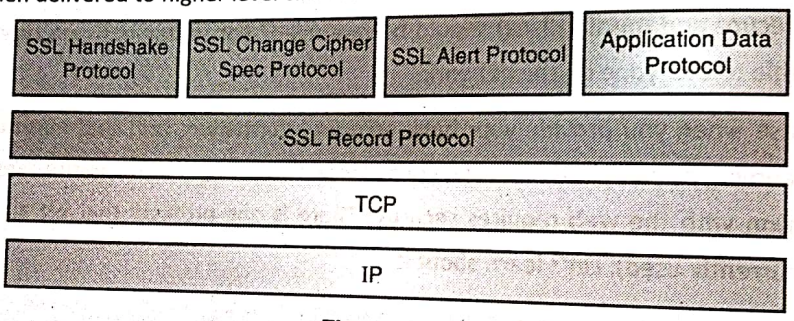


Fig. 4.4.1

- Let's learn about each one of them in detail.

4.4.1(A) Session and Connection States

- An SSL session is stateful which means that parameters negotiated during the session establishment persist (stay the same) until the session is terminated. The SSL handshake protocol coordinates the states of the client and server. It is thus important to preserve Session and Connection States.
- Table 4.4.2 summarizes a Session State.

Table 4.4.2 : Session state

Sr. No.	Fields	Purpose
1.	Session Identifier	A session ID chosen by the server to identify an active or resumable session state.
2.	Peer Certificate	X.509 Certificate of the other party in the communication.
3.	Compression method	The algorithm used to compress data prior to encryption.
4.	Cipher Specification	Chosen encryption algorithm such as AES and a hash algorithm such as SHA.

Sr. No.	Fields	Purpose
5.	Master Secret	48-byte secret shared between the client and server.
6.	Is resumable	A Boolean flag indicating whether a session ID can be used to initiate new connections.

Table 4.4.3 summarizes a Connection State.

Table 4.4.3 : Connection state

Sr. No.	Fields	Purpose
1.	Server and client random	Byte sequences for establishing connection
2.	Server write MAC secret	The secret used in MAC operations on data written by the server
3.	Client write MAC secret	The secret used in MAC operations on data written by the client
4.	Server write key	The bulk cipher key for data encrypted by the server and decrypted by the client
5.	Client write key	The bulk cipher key for data encrypted by the client and decrypted by the server
6.	Initialization vectors	Random number to initialize encryption operation
7.	Sequence numbers	Sequence numbers for transmitted and received messages for each connection

4.4.1(B) SSL Record Layer Protocol

Definition : The SSL Record Layer is the last protocol that receives the raw data from the higher application layers and other SSL protocols such as handshake.

- Its core function is to facilitate (perform) data transfer. The basic unit of data in SSL is a record. Each record consists of a five-byte record header, followed by data.
- There are four types of records in SSL.

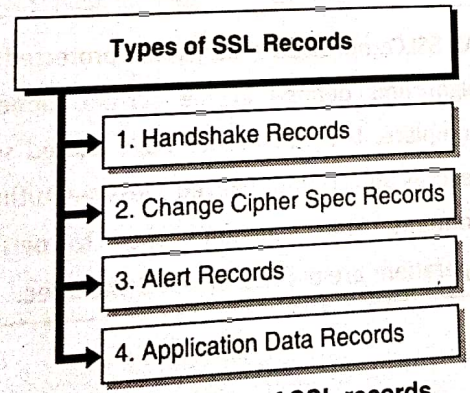


Fig. 4.4.2 : Types of SSL records

- The five-byte format of an SSL Record Header is as follows:

SSL record type (1-byte)	SSL Major Version (1-byte)	SSL Minor Version (1-byte)	Length of data in the record (2-bytes)

Fig. 4.4.3 : SSL record header format

- Fig. 4.4.4 shows a simplistic block diagram of steps involved in the SSL Record Protocol.

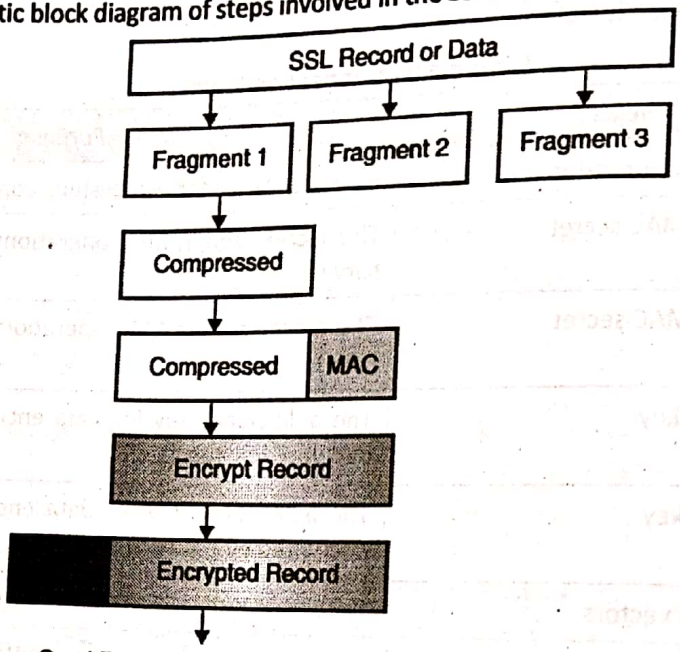


Fig. 4.4.4 : Block diagram of SSL record protocol

- At a high level the SSL Record Protocol performs three operations as shown in Table 4.4.4.

Table 4.4.4 : Operations of SSL record protocol

Sr. No.	Operation Performed	Purpose
1.	Fragmentation	Break original data into SSL Plaintext records of 2^{14} bytes or less
2.	Compression and Decompression	All SSLPlaintext records are compressed using the compression algorithm defined in the current session state. The compression algorithm translates an SSLPlaintext structure into an SSLCompressed structure
3.	Payload Protection	All SSLCompressed records are protected using the encryption and MAC algorithms defined in the current CipherSpec. Once the handshake is complete, the two parties have shared secrets that are used to encrypt records and compute keyed Message Authentication Codes (MACs) on their contents. The techniques used to perform the encryption and MAC operations are defined by the CipherSpec.

4.4.1(C) SSL Change Cipher Spec Protocol

Definition : *The Change Cipher Spec protocol notifies about the changes in cipher parameters.*

The protocol consists of a single message, which is encrypted and compressed. The Change Cipher Spec Protocol notifies the communicating parties about any change in the previously negotiated Cipher Specifications or Keys.

The keys or the algorithms need to be changed at times for reasons such as renewing the session or resuming the session. The change cipher spec message is sent by both the client and server to notify the receiving party that subsequent records will be protected under the just-negotiated CipherSpec and keys.

4.4.1(D) SSL Alert Protocol

Definition : *The SSL Alert Protocol signals problems with an SSL session.*

One of the content types supported by the SSL record layer is the alert type.

Alert messages notify the

i. Severity of the alert and

ii. A description of the alert

Alert messages with a severity level of *fatal* result in the immediate termination of the connection. In this case, other connections corresponding to the session may continue, but the session identifier is invalidated, preventing the failed session from being used to establish new connections. Like other messages, alert messages are encrypted and compressed, as specified by the current connection state.

Table 4.4.5 summarizes the various alert records.

Table 4.4.5 : Various alert records

Alert Code	Alert Message	Alert Level	Alert Description
0	close_notify	1 (Warning)	notifies the recipient that the sender will not send any more messages on this connection
10	unexpected_message	2 (Fatal)	An inappropriate message was received
20	bad_record_mac	2 (Fatal)	A record is received with an incorrect MAC
30	decompression_failure	2 (Fatal)	The decompression function received improper input
40	handshake_failure	2 (Fatal)	The sender was unable to negotiate an acceptable set of security parameters given the options available
41	no_certificate	1 (Warning)	Sent in response to a certification request if no appropriate certificate is available
42	bad_certificate	1 (Warning)	A certificate was corrupt
43	unsupported_certificate	1 (Warning)	A certificate was of an unsupported type
44	certificate_revoked	1 (Warning)	A certificate was revoked by its signer
45	certificate_expired	1 (Warning)	A certificate has expired or is not currently valid
46	certificate_unknown	1 (Warning)	Some other (unspecified) issues
47	illegal_parameter	2 (Fatal)	A field in the handshake was out of range or inconsistent with other fields

The alert record consists of 2 bytes of information from Fig. 4.4.5.

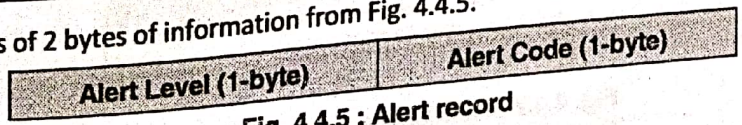


Fig. 4.4.5 : Alert record

4.4.1(E) SSL Handshake Protocols

SPPU - May 19

Q. Explain secure socket layer handshake protocol in detail. **(May 19, 7 Marks)**

Definition : *The cryptographic parameters of the session state are produced by the SSL handshake protocol.*

- When an SSL client and the server first start communicating, they need to agree upon certain parameters. There are also several steps that need to be carried out to establish a secure session. At a high level, the following four steps are carried out.

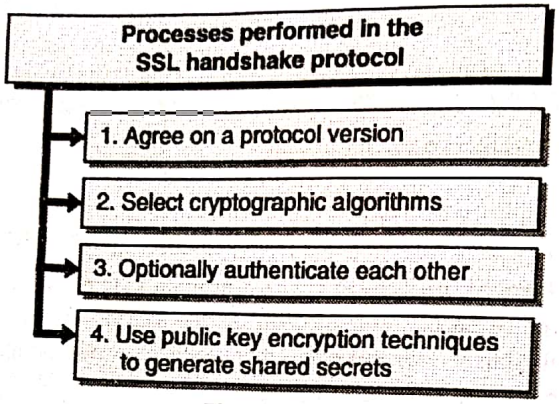


Fig. 4.4.6

- Fig. 4.4.7 illustrates the detail steps of handshake process diagram.

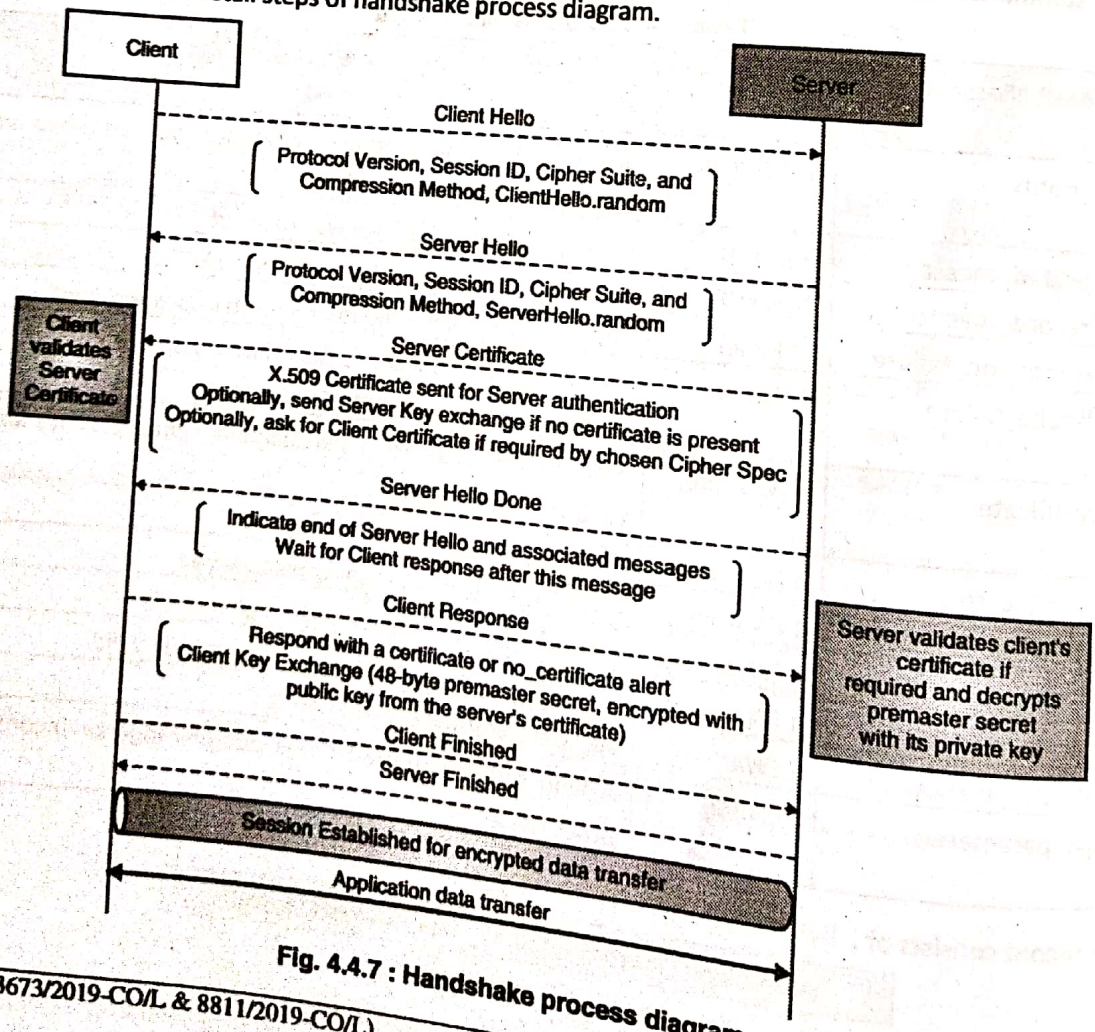


Fig. 4.4.7 : Handshake process diagram

Step 1 : Hello messages (Establish security capabilities)

The hello phase messages are used to exchange security enhancement capabilities between the client and server.

1. **Client Hello**

When a client first connects to a server it is required to send the client hello as its first message. The client can also send a client hello in response to a hello request or on its own initiative in order to renegotiate the security parameters in an existing connection. The list of parameters sent is in Fig. 4.4.7.

2. **Server Hello**

The server processes the client hello message and responds with server hello message. The list of parameters sent is in Fig. 4.4.7.

Step 2 : Server Authentication and Key Exchange

- This is the most important step. This is why you need SSL at all. Before proceeding to interact with the server you should find out "Is this really the server you want to talk to?" This is crucial. For example, if you want to do a banking transaction, before providing your account information, username and password, you MUST validate that the website you are on (server behind the website) is legitimate.
- In this step, the client validates the server certificate. Any certificate related errors are highlighted.

Step 3 : Client authentication and Key exchange

If the certificate is found valid, client exchanges the keying material that would be subsequently used to encrypt the messages. The client generates a 48-byte premaster secret, encrypts it using the public key from the server's certificate and sends the result in an encrypted pre-master secret message.

Step 4 : Connection establishment and data transfer

Once all the connection parameters are negotiated and exchanged, a connection between the server and the client is established. Once the connection is established, the data transfer begins between the server and the client. The data is encrypted based on the negotiated terms.

4.4.2 Transport Layer Security (TLS)

- As you learnt earlier, SSL is obsolete. TLS replaced SSL in 1999. The underlying working of TLS is very similar to SSL.
- TLS is more efficient and secure than SSL. It provides stronger message authentication, key-material generation and supports pre-shared keys, secure remote passwords, elliptical-curve keys and Kerberos. TLS and SSL are not interoperable, but TLS provides backward compatibility for devices using SSL.

4.5 HTTPS

Now that you learnt how SSL works, let's learn about one of its implementations – HTTPS application protocol.

Definition : HTTPS establishes a secure SSL/TLS tunnel before beginning data transfer.

- The Hypertext Transfer Protocol (HTTP) is an application protocol used to transfer data on distributed and connected systems. HTTPS is the secure version of HTTP. The 'S' at the end of HTTPS stands for 'Secure'. It means that all the communications between your client (browser, mobile apps) and the server (website, web application) is encrypted. HTTPS is often used to protect confidential online interactions such as online banking.
- Conceptually, HTTPS is very simple. Simply use HTTP over TLS (previously SSL) instead of HTTP. The use of TLS (previously SSL) ensures that the adequate protection mechanisms such as encryption, server authentication, hashing, and optionally client authentication are effectively applied, and the communication is adequately protected.

4.5.1 Comparison between HTTP and HTTPS

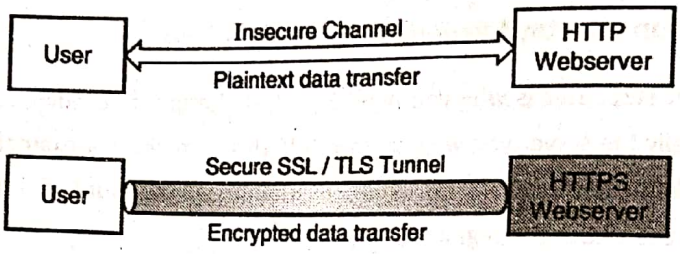


Fig. 4.5.1

Sr. No.	HTTP	HTTPS
1.	Data transfer in plaintext	Data transfer in ciphertext
2.	Default port is 80	Default port is 443
3.	Does not require SSL/TLS or Certificates	Requires SSL/TLS implementation with Certificates
4.	URL has http://	URL has https://
5.	Should be avoided	Should be preferred
6.	Search engines do not favour the insecure websites	Improved reputation of the website in search engine
7.	Users worried about their data	Users confident about the security of their data

4.5.2 Motivation / Benefits of using HTTPS

1. Increasing sensitivity of data

With the proliferation (widespread use) of internet, a lot of sensitive communication such as online banking, ticketing, shopping, etc. is taking place over the Internet. There is an ever increasing need to ensure that the communication is secure (confidentiality and integrity of the information is enforced).

Information such as your password or credit card number is not transferred in plaintext that can potentially be captured and then misused.

2. **Authentication**

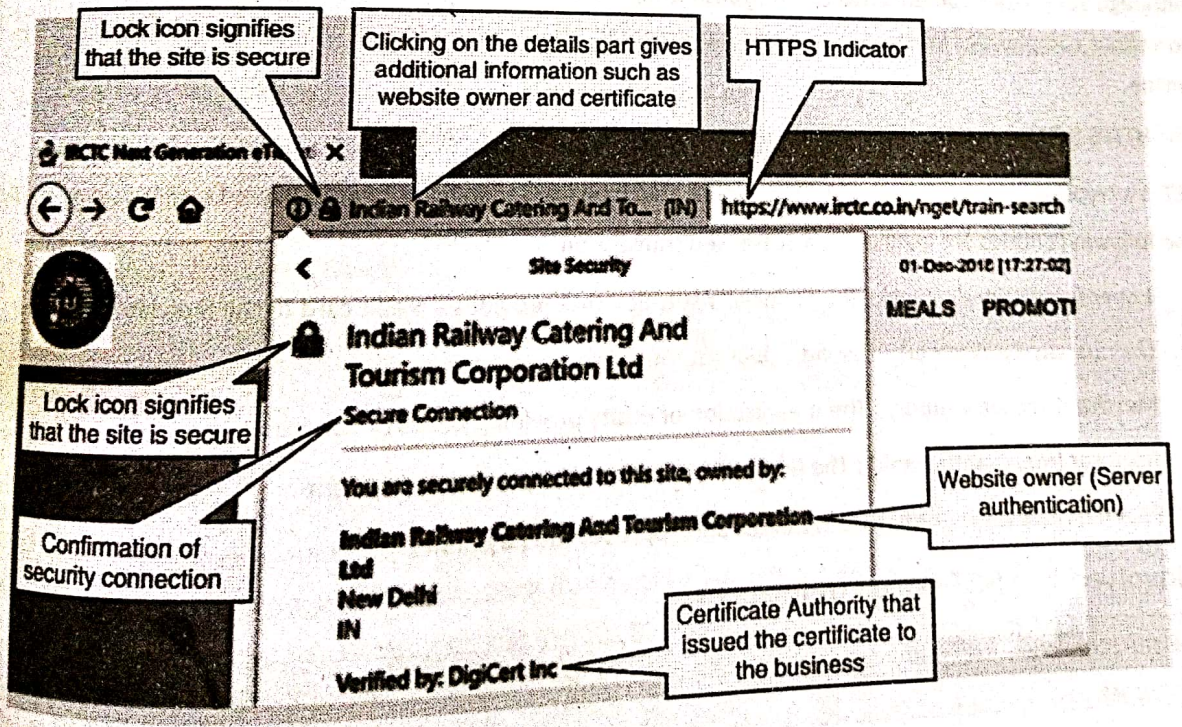
One of the critical use cases that HTTPS serves is that using it you can potentially authenticate a website or a business. HTTPS connection is established using X.509 certificates and certificate authorities do proper business or website validation before issuing certificates. Certificates help you prove that a website is indeed legitimate, and you are indeed interacting with the right website. This avoids several online frauds where a similar looking banking or e-commerce site can capture your confidential details.

3. **Privacy requirements**

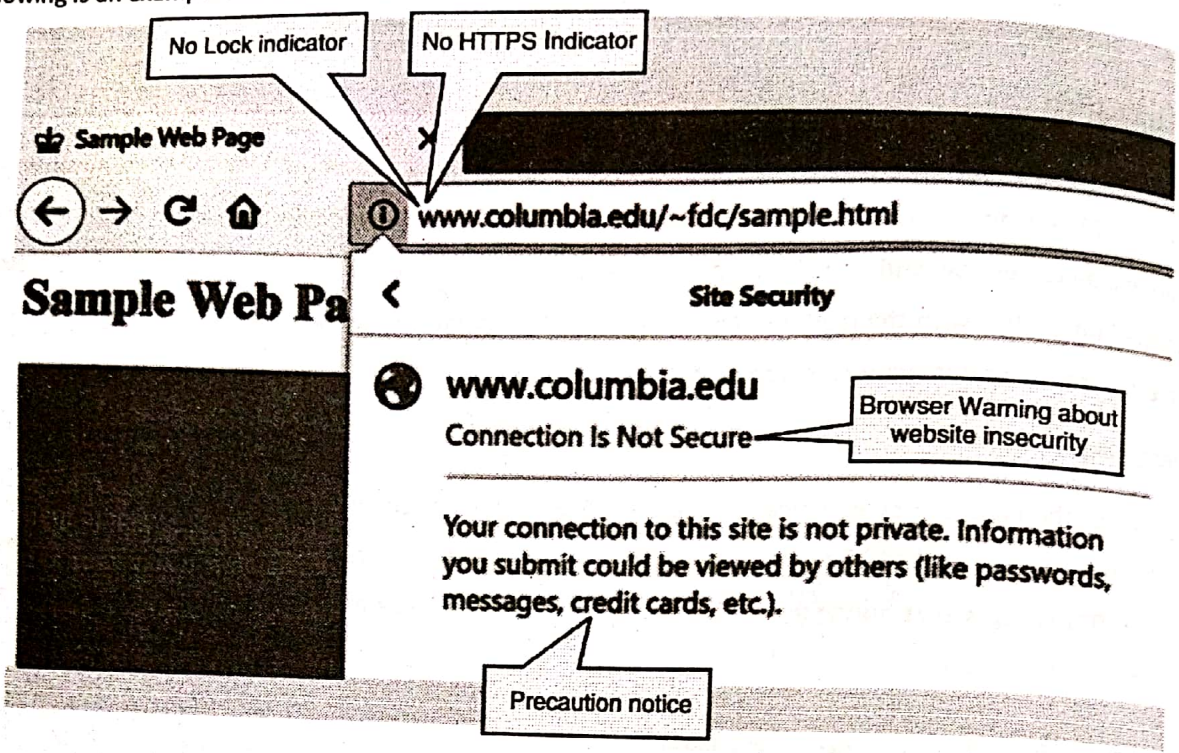
Often times, the nature of communication is private even if it is not confidential. For example, your health reports, your chats, your location details, etc. require that they are adequately protected when transferred over the network. Use of HTTPS ensures that encryption is applied to all data seamlessly and the private information is adequately protected during transfer.

4.5.3 Format, Port Number and Representation

- Typical format of HTTPS is https://www.example.com. It works over port 443 by default. You would have seen various websites with HTTPS enabled. These days browsers show green color in the URL for HTTPS protected websites and warning for non-HTTPS websites.
- Following is an example of a HTTPS protected website.



– Following is an example of a non-HTTPS website.



4.6 Secure Electronic Transactions (SET)

Definition : *Secure Electronic Transaction (SET) is a security technology proposed by Visa and MasterCard to allow for secure credit card transaction.*

– Although SET provides an effective way of transmitting credit card information, it is not used in the industry today due to various complexities involved. The world has accepted SSL as the preferred way of conducting secure transaction online.

– The HTTPS (HTTP over SSL) provides a secure way to carry out online transactions.

– SET is a cryptographic protocol and infrastructure developed to send encrypted credit card numbers over the Internet. The following entities are involved in a SET-based transaction.

1. **Issuer (cardholder's bank) :** The financial institution that provides a credit card to the individual
2. **User (or the Cardholder) :** The individual authorised to use a credit card
3. **Merchant (or the vendor) :** The organisation or entity providing goods or services
4. **Acquirer (merchant's bank) :** The financial institution that processes payment
5. **Payment gateway :** Entity that initializes and approves the payment

Steps involved in a SET-based transaction

Note : Both cardholders and merchants must register with CA (certificate authority) first, before they can buy or sell on the Internet.

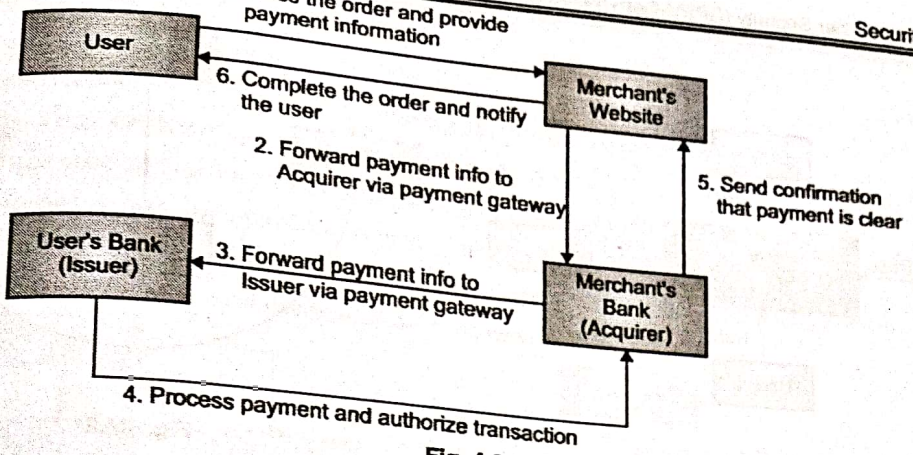


Fig. 4.6.1

1. User browses the merchant's website and decides the order. She sends the order and payment information.
   a. **Purchase Order :** This part is for merchant.
   b. **Card Information :** This part is for merchant's bank only (acquirer).
2. Merchant forwards the card information to its bank via the payment gateway.
3. Merchant's bank sends the credit card information to the Issuer via the payment gateway.
4. Issuer verifies the credit card information and sends the authorisation to the Merchant's bank.
5. Merchant's bank send authorisation to merchant notifying that the payment is cleared.
6. Merchant completes the order and sends confirmation to the customer.

4.7 Email Security

- Around billions of emails are sent across the globe every day. Emails have become the primary source of official communication. With such a wide use of emails, attackers are inclined and motivated to intercept emails and get the message and at times modifying the messages before the recipient gets it.
- It is important that you secure the email communication as any other form of communication. In this section, you will learn about a couple of email security standards that you could use.

4.7.1 Pretty Good Privacy (PGP)

***Definition :** Pretty Good Privacy (PGP) is an email security program that was developed in 1991. It is based on public key cryptography.*

4.7.1(A) Web of Trust

- In Public Key Cryptography system that depends upon a third-party Certificate Authorities (CAs) to establish trust, there is no mutual trust amongst the users. Each user trusts a reputed CA and thus CA plays a predominant role in establishing the trust so that communication can happen between users. If there is no CA, the trust relationship is not established and thus the communication may not happen.

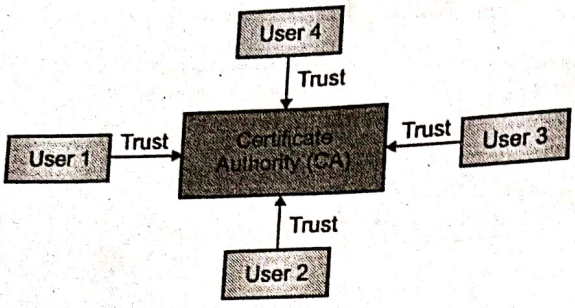


Fig. 4.7.1

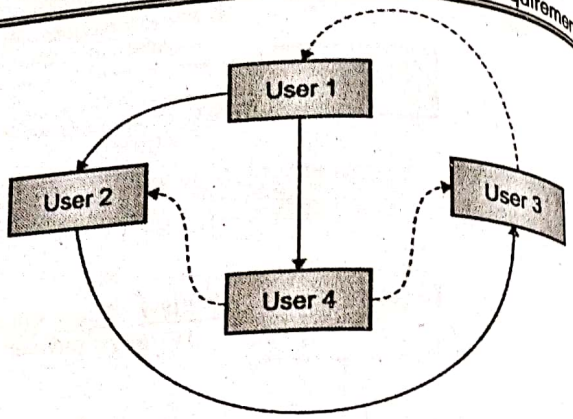


Fig. 4.7.2

- Unlike the traditional Public Key Cryptography system that depends upon a CA to establish trust amongst the users, the earlier implementations of PGP did not use the regular CAs for issuing certificates. It used "Web of Trust" where each user generates and distributes his or her public key, and users sign each other's public keys, which creates a community of users who trust each other. This is different from the CA approach, where no one trusts each other; they only trust the CA.
- So, basically, PGP is a system of "I don't know you, but my friend Alice says that you can be trusted, so I will trust you on her words". In the figure, as you understand, there is a trust relationship (User 1, User 2) and (User 2, User 3). Now, when User 3 needs to communicate with User 1, it establishes a trust inherited from its prior trust on User 2.
- There is no third-party involved in this scenario. Each user keeps in a file, referred to as a key ring, a collection of public keys he has received from other users. Each key in that ring has a parameter that indicates the level of trust assigned to that user and the validity of that particular key.

4.7.1(B) PGP Services

PGP provides the following services. You can use one or more services at a time. For example, if you intend to use only encryption service, you can do so. If you intend to use only the digital signature service, you can do so. Let's learn a brief about these services.

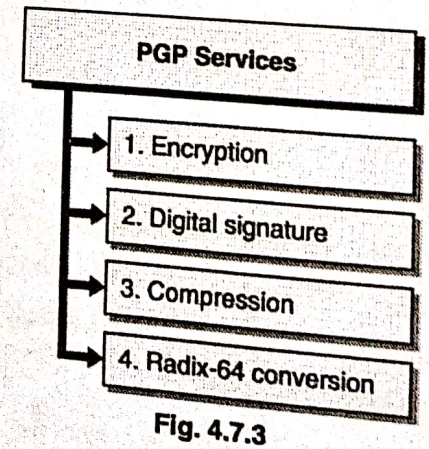


Fig. 4.7.3

1. **Encryption**

a. The sender creates a message.

b. PGP generates a random number that is used as symmetric key to encrypt it.

c. The symmetric key is encrypted using receiver's public key.

d. Encrypted message and the encrypted symmetric key are sent to the receiver.

e. The receiver decrypts the encrypted symmetric key using her private key.

f. Once the receiver gets the symmetric key after decryption, the key can be used to decrypt the message.

Fig. 4.7.4 is a simplistic diagram of encryption steps followed by PGP.

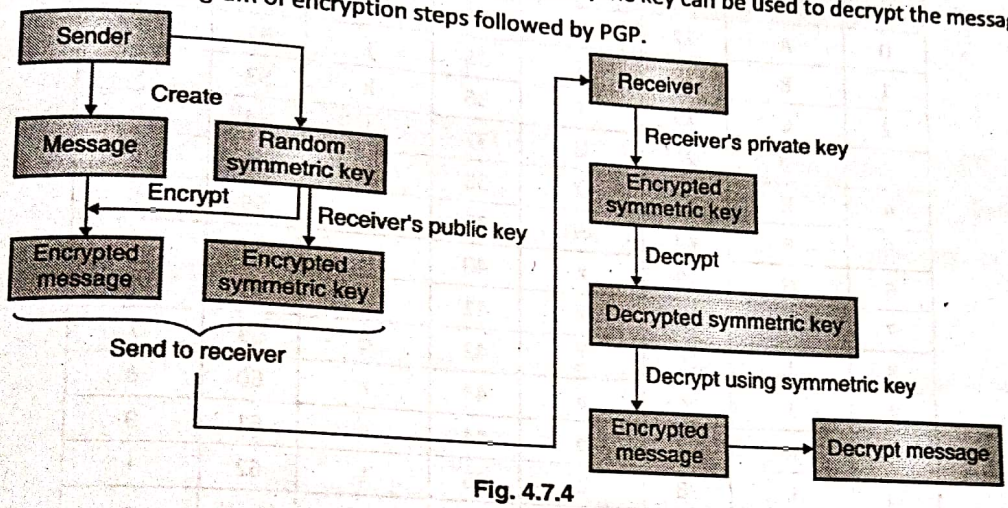


Fig. 4.7.4

2. **Digital Signature**

The digital signature uses a hash code or message digest algorithm, and a public-key signature algorithm. You have already learnt digital signature in detail in Unit 2. Refer it for a quick refresher.

3. **Compression**

PGP compresses the message after applying the signature but before encryption. Compression has the added side effect that some types of attacks can be avoided by the fact that even the slightly altered, compressed data does not decompress without errors. This side security benefit is operationally useful.

4. **Radix-64 conversion**

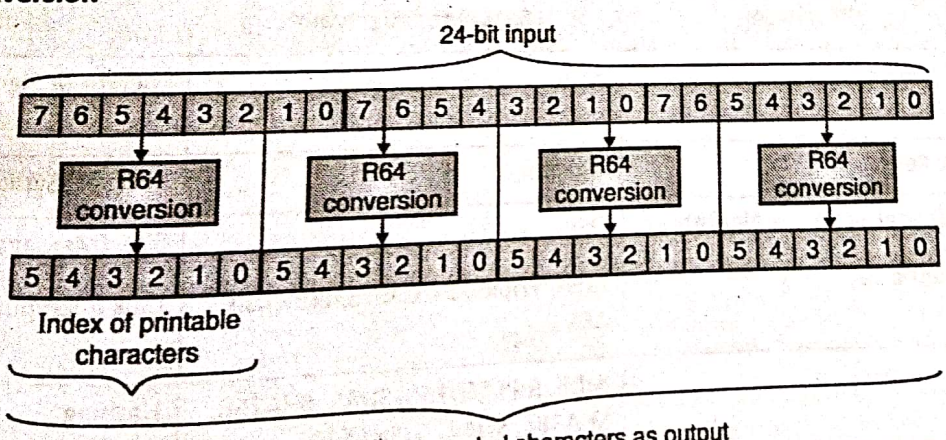


Fig. 4.7.5

R64 conversion is useful for compatibility of emails across varied systems. PGP's underlying native representation for encrypted messages, signature certificates, and keys is a stream of arbitrary octets. Some systems only permit the use of blocks consisting of seven-bit, printable text. So, for transporting PGP's native raw binary octets through channels that are not safe to raw binary data, a printable encoding of these binary octets is needed. PGP provides the service of converting the raw 8-bit binary octet stream to a stream of printable ASCII characters, called Radix-64 encoding.

Each 6-bit group is used as an index into an array of 64 printable characters as shown in Table 4.7.1.

Table 4.7.1 : Encoding map

Value	Encoding	Value	Encoding	Value	Encoding	Value	Encoding
0	A	17	R	34	i	51	Z
1	B	18	S	35	j	52	0
2	C	19	T	36	k	53	1
3	D	20	U	37	l	54	2
4	E	21	V	38	m	55	3
5	F	22	W	39	n	56	4
6	G	23	X	40	o	57	5
7	H	24	Y	41	p	58	6
8	I	25	Z	42	q	59	7
9	J	26	a	43	r	60	8
10	K	27	b	44	s	61	9
11	L	28	c	45	t	62	+
13	N	30	e	47	v	(pad)	=
14	O	31	f	48	w		
15	P	32	g	49	x		
16	Q	33	h	50	y		

4.7.1(C) PGP Algorithms

Table 4.7.2 is a summary of various PGP services and algorithms they support.

Table 4.7.2 : Summary of various PGP services and algorithms

Sr. No.	PGP Service	Supported Algorithm	Purpose
1.	Public Key	RSA	Encrypt or Sign (Symmetric Key)
2.	Public Key	Elgamal	Encrypt (Symmetric Key)
3.	DSA (Digital Signature Algorithm)	DSS	Sign (Message)
4.	Symmetric Key	IDEA, TripleDES, CAST5, Blowfish, AES	Bulk encryption (message)
5.	Hash	MD5, RIPEMD160, SHA1, SHA256, SHA384, SHA224, SHA512	Hashing
6.	Compression	ZIP, ZLIB, BZip2	Compress messages

4.7.2 MIME

- MIME is an acronym that stands for Multipurpose Internet Mail Extensions.

Definition : *MIME (Multi-Purpose Internet Mail Extensions) is an extension of the original Internet e-mail protocol that lets users to use the protocol to exchange different kinds of data files (audio, video, images, application programs, and others) via email.*

- Today, you can, thus, use email for attaching various kind of files and send it across. MIME does not provide security specifications for sending and receiving emails securely. Hence, S/MIME protocol is used.

4.7.3 S/MIME

Definition : *S/MIME (Secure/Multipurpose Internet Mail Extensions) provides a consistent way to send and receive secure MIME data (emails).*

It is based on certificates (Public Key Cryptography) and works as you have learnt in previous sections and units.

4.7.3(A) S/MIME Services

- S/MIME provides the following cryptographic security services for electronic messaging applications as shown in Fig. 4.7.6.

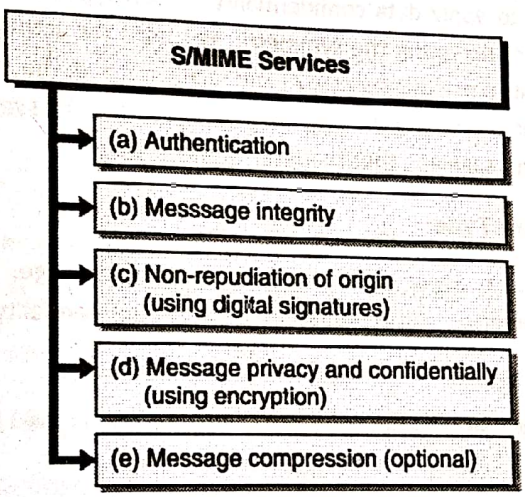


Fig. 4.7.6 : S/MIME Services

- These services are provided using same techniques as you have learnt in the previous sections and units.

4.7.3(B) S/MIME Algorithms

Table 4.7.3 is a summary of various S/MIME services and algorithms they support.

Table 4.7.3 : S/MIME supported services and algorithms

Sr. No.	S/MIME Service	Supported Algorithms	Purpose
1.	Message Integrity	SHA-256, SHA-1, MD5	Hashing
2.	Non-repudiation, Authentication	RSA and DSA with Hashing algorithms	Digital Signature
3.	Key Encryption	RSA, RSAES-OAEP, Diffie-Hellman	Encrypting Symmetric key
4.	Privacy and Confidentiality	AES, DES, Triple DES	Message Encryption

4.7.3(C) S/MIME Cryptographic Message Syntax (CMS)

- S/MIME standard describes a protocol for adding cryptographic signature and encryption services to MIME (email) data. The MIME standard provides a general structure for the content of Internet messages and allows extensions for new content-type-based applications.
- The S/MIME specification defines how to create a MIME (email) body part that has been cryptographically enhanced according to the Cryptographic Message Syntax (CMS).
- There are 4 types of CMS used in S/MIME:

1. Data Content Type

This is the original plaintext form of the email message. It has an identifier that is referred whenever it is compressed, encrypted or digitally signed.

2. SignedData Content Type

SignedData content type is used when a sender needs to apply a digital signature to a message. Applying a signature to a message provides authentication, message integrity, and non-repudiation of origin.

3. EnvelopedData Content Type

This content type is used to apply data confidentiality (via encryption) to a message. A sender needs to have access to a public key (for encrypting the symmetric key used for actual encryption of the message) for each intended message recipient to use this service. At the receiver's end, the receiver uses her private key to decrypt the symmetric key used for encrypting the original message. Once the symmetric key is available, it can be used to decrypt the encrypted message and thus read the email message.

4. CompressedData Content Type

This content type is used to apply data compression to a message. This content type does not provide authentication, message integrity, non-repudiation, or data confidentiality. It is only used to reduce the size of the message.

4.7.3(D) Comparison between PGP and S/MIME

Q. Compare PGP, MIME and S/MIME. SPPU – May 19 (May 19, 6 Marks)

Sr. No.	Comparison Attribute	PGP	S/MIME
1.	Trust established using	Web of Trust	Public Key Infrastructure
2.	Provides Authentication	No	Yes
3.	Used for	Securing text messages only	Securing Messages and attachments
4.	Industry use	Less Common	Widely used
5.	Administrative overhead	High	Low
6.	Cost	High	Low
7.	Convenience	Low	High

CHAPTER 5 UNIT V

Firewall and Intrusion

Syllabus :

At the end of this unit, you should be able to understand and comprehend the following syllabus topics :

- Firewall
  - Introduction
  - Characteristics and types
  - Benefits and limitations
  - Firewall architecture
- Computer Intrusions and Intrusion detection systems (IDS)
  - Need
  - Methods
  - Types of IDS
- Access Control
- Trusted Systems
- Password Management, Limitations and Challenges

5.1 Firewalls

- Your computer is connected to the internet. How do you protect it from someone trying to access it over the internet? How do you prevent some rogue programs on your computer to send information to the attacker? Firewalls could be a mechanism.

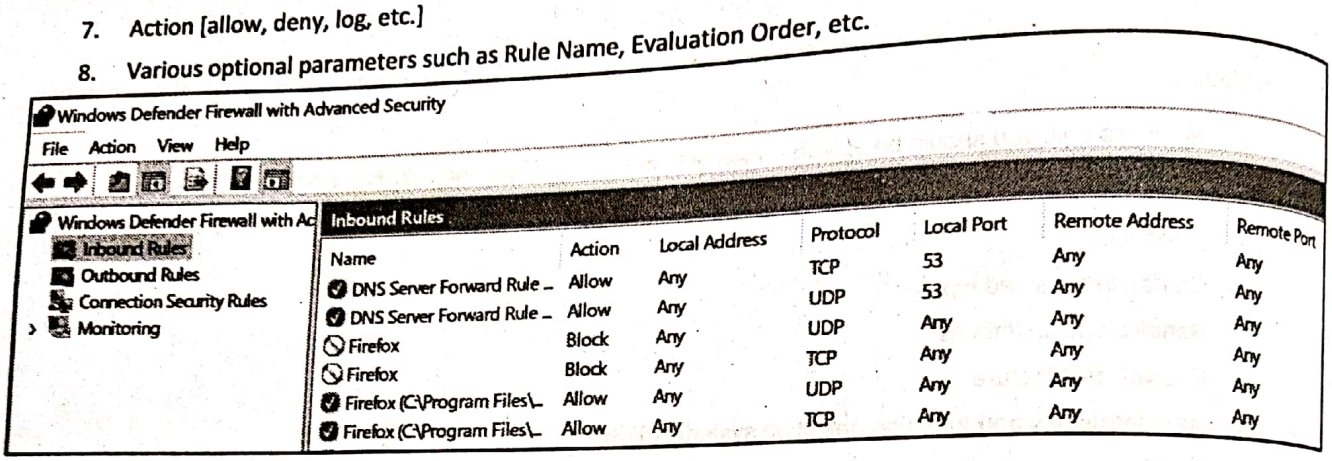
Definition : *Firewalls are network security systems that protect the computing resources on a trusted network from unauthorized access.*

- For example, you can access google.com website but not it's webserver's operating system. If you try connecting to the webserver except over the HTTP or HTTPS, the connection would be denied. That's what a firewall does at a high level.
- You need to define various rules, as per your security requirements, in the firewall and the firewall evaluates those rules before granting or denying access to the requested resource.

Components of a firewall rule

- Typically, a firewall rule consists of the following parameters :
  1. Source IP address or hostname

2. Destination IP address or hostname
3. Source Port number
4. Destination Port number
5. Direction of communication [inbound or outbound]
6. Protocol name [TCP, UDP, ICMP or various others]
7. Action [allow, deny, log, etc.]
8. Various optional parameters such as Rule Name, Evaluation Order, etc.



– This is a snapshot of Microsoft® Windows® Firewall.

5.1.1 Classification of Firewalls

SPPU – May 19

Q. What are the various types of firewall ? **(May 19, 4 Marks)**

Q. Explain packet filtering firewall. **(May 19, 4 Marks)**

Firewalls can be classified based on various attributes. Fig. 5.1.1 shows types of firewalls.

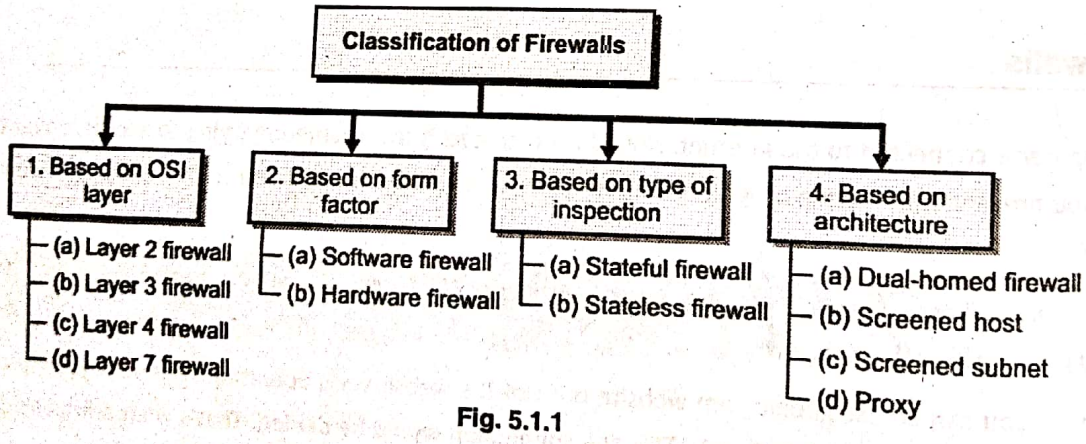


Fig. 5.1.1

1. Based on the OSI Layer

As you understand, OSI is a conceptual networking model. Based on the various layers, firewalls can be classified as following:

a. **Layer 2 Firewall :** These firewalls work at the "Data Link" layer of the OSI model. These firewalls require MAC, VLAN or device hardware level information to operate. One of the greatest advantage of these types of firewalls is that they are not IP dependent.

b. **Layer 3 Firewall :** These firewalls work at the "Network" layer of the OSI model. These filter traffic based on source/destination IP, port, and protocol. These are one of the most prevalent types of firewalls in use today. These are also called as Stateless firewalls. These are also called *first-generation* firewalls.

c. **Layer 4 Firewall :** These firewalls work at the "Transport" layer of the OSI model. These firewalls do everything that a Layer 3 firewall does and additionally track the active network connections and allow/deny traffic based on the state of those connections.

These can effectively stop DoS attacks such as the ones based on TCP SYN/ACK as these are aware of the state of connection. These are also called as Stateful firewalls. These are also called *second-generation* firewalls.

d. **Layer 7 Firewall :** These firewalls are called Layer 7 but can work at three layers – Session, Presentation and Application. For simplicity, these are just called Layer 7 firewalls. Layer 7 firewalls do everything that a Layer 4 firewall does and additionally include the ability to intelligently inspect the contents of the network packets passing through them. For example, a Layer 7 firewall could deny all the HTTP requests from Korean IP addresses. They have the actual packet content level visibility and are the most advanced types of firewall in use today. These are also called *third-generation* firewalls.

2. Based on the form factor

Form factor or the footprint is the way the firewall is actually packaged and deployed. They can be classified as,

a. **Software Firewalls :** These firewalls work as a software program and require an operating system to run them. All the implementation logic is coded in software and they are installed, patched, upgraded and maintained like a regular computer software. These firewalls could work at any of the OSI layers as discussed before.

b. **Hardware Firewalls :** Firewalls can also be deployed as a hardware device. Hardware firewall may have better performance and they come packaged in a ready to use hardware device. Like any other firewall, you need to configure it as per your security requirements.

1. Based on the type of inspection

Firewalls can keep track of connections or just work based on the configured rules. Based on their inspection types, these can be classified as

a. **Stateful Firewalls :** These firewalls keep track of the state of connections apart from the defined firewall rules. These precisely understand various handshake protocols and can effectively stop attacks that try to manipulate connection establishment or maintenance process.

b. **Stateless Firewalls :** Stateless firewalls typically work at the Layer 3 and take decisions based on the defined rule parameters such as IP, Port and Protocol. These do no track connection states and cannot effectively protect against attacks that manipulate connection processes.

4. Based on architecture

Firewalls can be deployed in many ways. They have special properties that make them suitable for one deployment type over the another. Based on the deployment possibilities, firewalls can be classified as

a. **Dual-homed Firewalls :** A Dual-Homed Firewall has two interfaces – one facing the external network and the other facing the internal network. It receives the external packets on one of its interfaces, evaluates firewall rules, and passes on the traffic to the designated internal resources via the second interface. The two interfaces are kept separate to isolate the external traffic with the internal traffic physically.

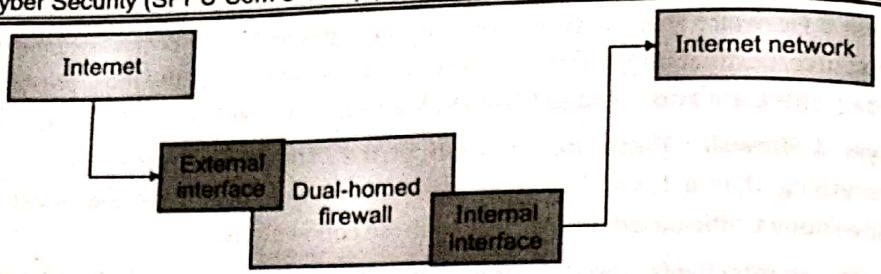


Fig. 5.1.2

b. **Screened Host :** In a screened host firewall, all internet (and other regulated) traffic goes through the firewall, no matter what. The internet router device first screens (filters) all the packets that are relevant to the network and then passes it to the Screened Host firewall for further inspection and applying rules.

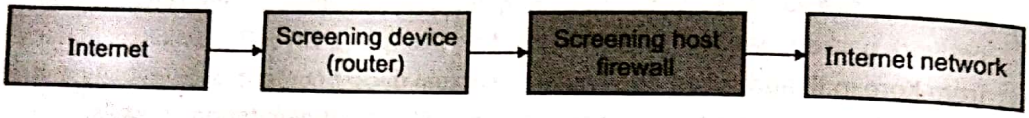


Fig. 5.1.3

c. **Screened Subnet :** In screened subnet architecture, two firewalls are used. One just after the external network and the one just before the internal network. Any network that lies between the two firewalls is called a Demilitarized Zone (DMZ). You place your public facing servers such as webservers, email servers etc. in DMZ. An attacker would have to bypass both the firewalls before she can hit the internal network. This kind of architecture is commonly used in the industry today.

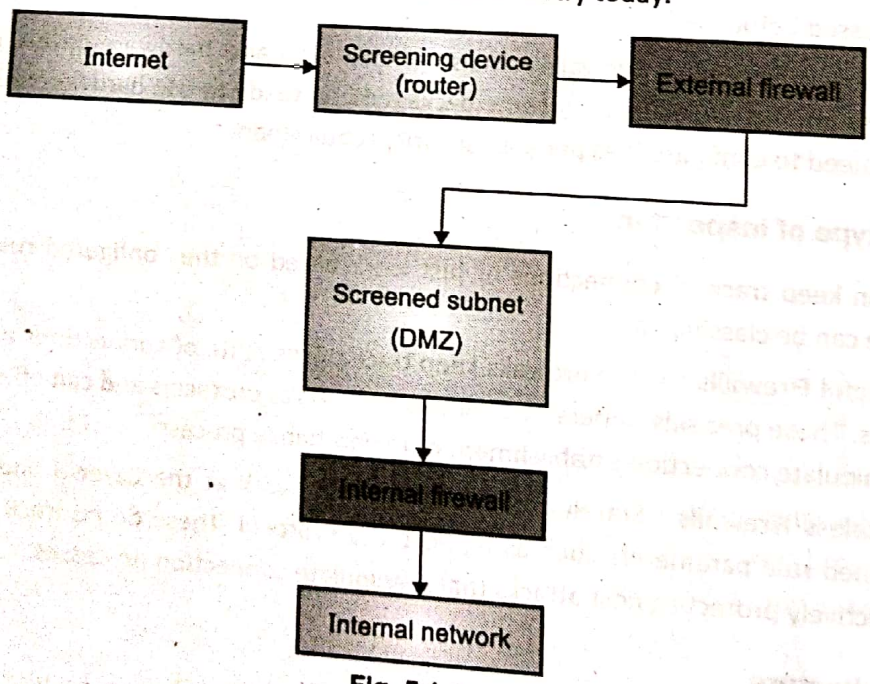


Fig. 5.1.4

d. **Proxy :** A proxy firewall stands between the trusted and the untrusted network and takes allow or deny decisions after careful inspection of what is being passed along. Like a regular proxy, the proxy firewall breaks the connection between the source and the destination. After examining the traffic, it self-establishes a connection with the destination and passes the intended traffic to the destination as if the packets were originating from it.

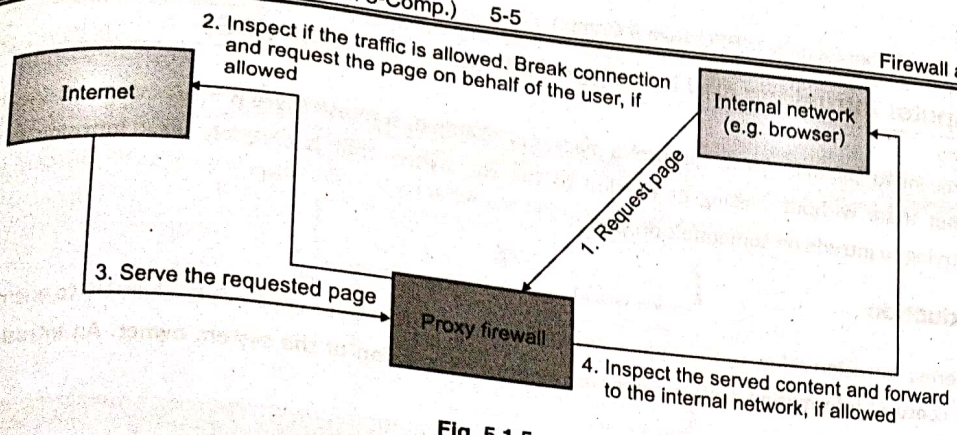


Fig. 5.1.5

5.1.2 Challenges in Managing and Deploying Firewalls

Q. Discuss limitations of firewall. **SPPU – May 19**

(May 19, 4 Marks)

Note : Irrespective of what challenges or limitations firewalls may have, they are heavily used throughout the industry. You cannot just imagine any network without several firewalls in place to monitor, inspect and manage legitimate traffic and separate it from the illegitimate traffic. So, just know what some of the management or technical challenges are and do not worry too much about them.

1. **Performance**

Since the traffic (as well as the content) needs to pass through the firewalls, there is a little performance degradation of the network. The adequate traffic examination may add up to a few milliseconds of latency on each packet.

2. **Business agility**

Firewall rules are usually manually added, edited or deleted. The pace of business might be too high to require several changes to the firewall rules frequently. Keeping up with these changes without making errors is difficult.

3. **Costs**

Modern (or advanced) firewalls that provide content and protocol level inspection may be cost-prohibitive for small or medium sized organizations.

4. **Insider attacks**

Firewalls are usually designed and deployed to protect a trusted network from an untrusted network. But, if there were other vulnerabilities (such as a missing OS security patch) that were exploited such that an attacker is already on the trusted network, firewalls might not be able to protect or limit damages to the other resources on the trusted network.

5. **Managing firewalls themselves**

Like your OS, printers or other software or hardware devices, firewalls need to be installed, patched, updated, etc. to remain operational. This adds a management overhead. Additionally, firewalls could have known vulnerabilities that need to be patched else a firewall that itself is lacking protection may not be very useful in providing you the required level of protection.

5.2 Computer Intrusions and Intrusion Detection Systems (IDS)

Intrusion means to encroach (or to capture) a place. For example, suppose there is a vacant site and you manage to build a small hut there without seeking permission of the site owner that is precisely what intrusion is. You, as an individual, are trying to intrude on someone's property.

5.2.1 Introduction

- In digital terms, intrusion refers to the similar situation where the malicious code or attackers try to encroach (forcibly enter and capture) information systems without requiring permission of the system owner. An Intrusion Detection System (IDS) is defined as,

Definition : A software that helps to find out if a system is breached.

Note : Breach is a word used in information security domain to describe any form of attack or unauthorized actions. You can use this word to mean anything that refers to the undesired actions and outcomes with respect to information security.

- So, in a nutshell, IDS can help you to find out if there were undesired actions or attacks carried out on your information systems. IDS works using various techniques as we will see later in this section. Note here that IDS does not help to prevent the attacks unlike anti-virus. It is only a system that can gather system information and find out if everything looks alright or not.

5.2.2 Need for IDS

SPPU - May 19

Q. Explain need of intrusion detection system. **(May 19, 2 Marks)**

IDS is one of the software-based security mechanisms that help to protect information system. At a high level, it is needed for the following reasons:

1. Defense in Depth

As you saw in the security architecture section, security is about minimizing the damage that can be possibly done. Defense in depth (or the layered approach) of security designing ensures that even if one of the controls is to fail, the overall security of the system would still be possibly healthy. IDS fulfill this need to bring an added layer of protection where any breaches or their possibilities can be identified quickly.

2. Automate intrusion detection

Imagine that you have a large set of machines, say 1,000 and more. How would you inspect each and every machine and find out if there were attacks or attempts to attack it? IDS helps you to automate this need and alert you when it detects any threat or likely a breach.

3. Corrective actions

Learning from threats or breaches that the IDS identifies, you can take corrective actions on your infrastructure design and could possibly strengthen its security. You might have some unprotected areas in your infrastructure that can be highlighted with the use of IDS.

5.2.3 Types of IDS

Q. Define signature based IDS. SPPU - May 19 (May 19, 3 Marks)

Broadly speaking, IDS can be classified based on what it monitors and how it monitors.

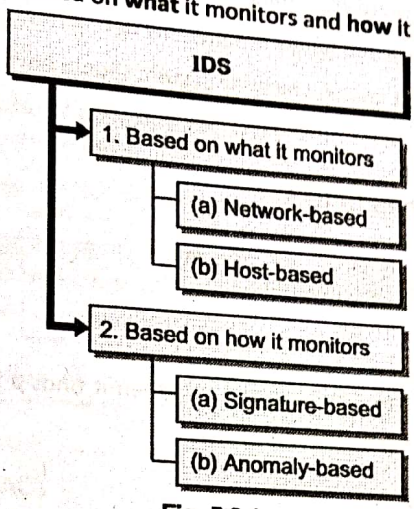


Fig. 5.2.1

1. Based on what it monitors

IDS can be classified into Network-based IDS (NIDS) and Host-based IDS (HIDS).

a. **Network-based IDS (NIDS) :** Network-based IDS evaluate intrusions from the networking side. They watch all network traffic as it reaches the various information systems. If there are any alerting situations based on the network traffic analysis, it notifies the administrator to take the corrective actions. NIDS do not have visibility into what's actually going on within the information system. It can only watch and detect threats and breaches from the networking viewpoint.

b. **Host-based IDS (HIDS) :** Host-based IDS are typically installed on the individual information systems and then they watch for suspicious activities occurring on the system. A system entity such as system services and processes, system files, privileged user actions, downloads etc. are closely monitored to detect any undesired activities. HIDS do not have visibility into what's going on at the networking side of the system. It can only watch and detect activities with respect to individual machines only.

2. Based on how it monitors

IDS can be classified into Signature-based and Anomaly-based.

a. **Signature-based :** Like banks and other organizations use human signature to validate requests and transactions, similarly Signature-based IDS has a pre-loaded database of various attack signatures (patterns of a possible attack). When it watches the activities, it constantly compares the activities' patterns with that in the database. If a match is found, it raises an alert. If you notice, there are 3 things to understand here:

- Signature based IDS can only detect attacks if it already and historically knows about an attack pattern.
- For new types of attack, signature-based IDS would not raise alerts.
- It is important for you to update the signature definitions time to time (like how you do in anti-virus system).

b. **Anomaly-based :** Anomaly typically means "deviation from routine". For example, if you wake up at 7 AM every day and one day you wake up at 4 AM that is an anomaly situation.

If I were to plot your wake-up time graph, 4 AM would show up away from your regular wake-up time. That 4 AM point on the graph is called outlier (or away from other samples). Similarly, the Anomaly-based IDS first establishes the baseline (common routine) of activities. It might take up to 2-3 weeks to "learn" what's right for a system. Once the learning phase is over, it would watch out for any activities that are not part of that baseline and raise alerts. If you notice, there are 3 things to understand here as well:

- It does not require signature and hence can possibly detect new attacks.
- It requires a learning period during which the system should have undergone all possible activities.
- If you plan to use the system for other purposes, you need to retrain the IDS.

5.2.4 Limitations and Challenges of IDS

SPPU – May 19

Q. Explain challenges of intrusion detection system. **(May 19, 3 Marks)**

1. Does not prevent attacks

As you understand, IDS can only detect and raise alerts when it finds a likelihood of a breach. It cannot prevent or block the breach from happening.

2. High rate of false alerts (noise)

IDS might generate a lot of false alerts. It could happen so for example, when there is a new traffic from a source that IDS has not seen before. You need to spend your resources to take a note of each alert and appropriately deal with it – either fix it or ignore it.

3. Complex systems

IDS systems are typically complex in nature and require regular administrative actions and tuning for adequate operations.

4. Bypassing IDS

Advanced attackers know what actions and activities a version and brand of IDS can detect and what not. They tune their activities to bypass such detection mechanisms and go undetected.

5.3 Access Control

Q. What is access control security services ? SPPU – May 19 **(May 19, 4 Marks)**

Definition : *Access Control Policies and Models dictate how and under what constraints (or conditions) principals (entities or subjects) access resources (objects).*

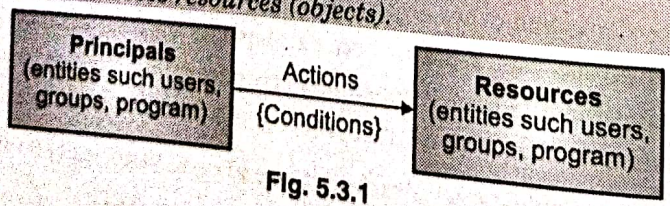


Fig. 5.3.1

- It follows the PARC model :
  - P = Principals (users, groups, programs)
  - A = Actions (Create, Read, Update, Delete)

- R = Resources (OS, Network, Files, etc.)
- C = Conditions (time of the day, type of OS, etc.)

- Access control is the primary way to restrict entities from
  - Interacting with unauthorised resources
  - Carrying out unauthorised actions on authorised resources
- There are 4 types of access control models.

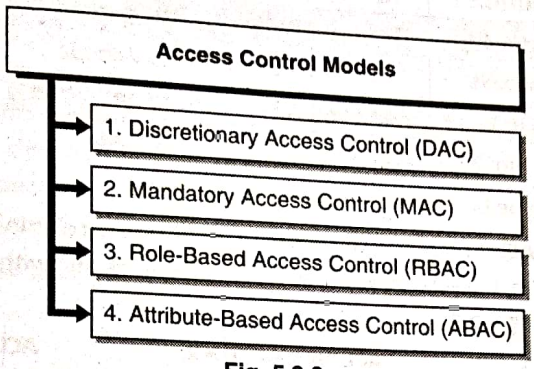


Fig. 5.3.2

1. Discretionary Access Control (DAC)

Definition : *Discretionary Access Control (DAC) is an access control model based on subject's identity and need to know.*

- This model is very simple to understand. As an owner of a particular resource, you can decide who uses that resource and with what permissions (how much). Access is provided based on individual's judgement on whom to provide access and how much.

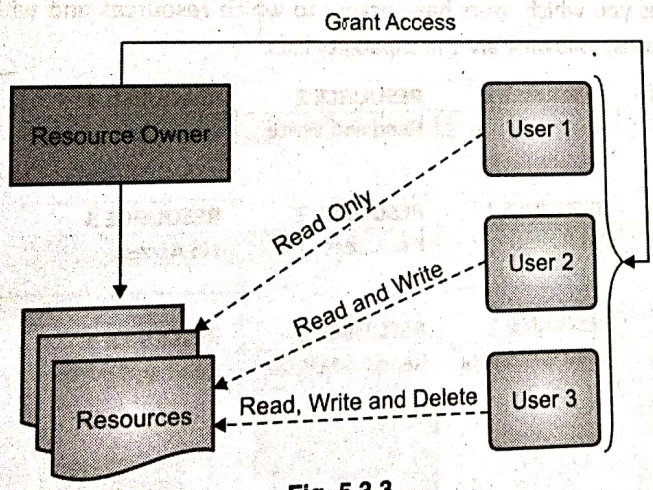


Fig. 5.3.3

- Let me give you an example. Suppose you own a car. You can hand over the keys to someone your trust who can drive your car outside the city limits. But, if you give your car for repairs, you would not like the car to be driven outside the city limits. The permissions assigned to the same resource (car) varies based on your judgement.
- Similarly, if you are an administrator on a system, you can decide which users (subjects or principals) to allow and what those users can do (which resources they can access and perform what actions on those resources).

- If you own certain files on the system, you can choose who else can access those files and in what capacity [read, write, delete, execute, share, print, etc.]. DAC is the most common type of access control model used. Typical operating systems such as Windows® and Linux enforce DAC model by default.
- DAC systems typically have an Access Control Matrix that lists the subjects and their rights (permissions) for respective objects (resources). Fig. 5.3.4 is a simplistic access control matrix.

USER	RESOURCE 1	RESOURCE 2	RESOURCE 3
JIM	Read only	Read and Write	No Access
JOHN	No Access	No Access	No Access
JACK	Read and Write	Read and Write	Read and Write

Fig. 5.3.4 : A simple access control matrix

- The Access Control Matrix can be broken and visualized as two parts.

1. **Access Control Lists (ACLs) :** Reading the access control matrix column-wise gives the access control list for a particular resource. It gives you who can use the resource and with what rights. For the given access control matrix, following are the ACLs.

ACL for Resource 1

USER	RESOURCE 1
JIM	Read only
JOHN	No Access
JACK	Read and Write

ACL for Resource 2

USER	RESOURCE 2
JIM	Read and Write
JOHN	No Access
JACK	Read and Write

ACL for Resource 3

USER	RESOURCE 3
JIM	No Access
JOHN	No Access
JACK	Read and Write

Access Control Lists

2. **Capability List :** Reading the access control matrix row-wise gives the capability of users across the resources. It gives you which users have access to which resources and with what rights. For the given access control matrix, following are the capability lists.

Capability of Jim

USER	RESOURCE 1	RESOURCE 2	RESOURCE 3
JIM	Read only	Read and Write	No Access

Capability of John

USER	RESOURCE 1	RESOURCE 2	RESOURCE 3
JOHN	No Access	No Access	No Access

Capability of Jack

USER	RESOURCE 1	RESOURCE 2	RESOURCE 3
JACK	Read and Write	Read and Write	Read and Write

Capability Lists

Advantages of DAC

1. Easy to understand.
2. Easy to implement.
3. Flexibility to assign / revoke the needed rights granularly (precisely) on the resources.

Disadvantages of DAC

1. It requires careful evaluation and administration of each resource, user and her rights periodically.
2. Once the permissions are granted, the user has unrestricted access to the resource until permissions are manually revoked.

3. If a user mistakenly installs a malware program, then the malware program automatically inherits the DAC permissions that were assigned to the user.
4. As the number of users increase, it becomes difficult to manage them and provide individual access right to each one of them.

2. Mandatory Access Controls (MAC)

Definition : Mandatory Access Control (MAC) is an access control model that restricts access to system resources based on their sensitivity and the authorisation level of the entity trying to access the resources.

- Systems that are designed with the MAC model have highly restrictive use of rights. Unlike DAC model, one user (or resource owner) cannot provide rights to another user. The access control decisions are automatically taken by the system based on the sensitivity of the resource and the authorisation level of the entity trying to access the resource. MAC models are highly secure and are predominantly used in the military environment and systems. One example of such a system used today is SELinux (Security Enhanced Linux.
- Let's dive a little deeper to understand the MAC model.

A. Security Labels

- It is crucial to understand security labels to make any sense of MAC model.

Definition : Security label is a way to associate a set of security attributes with a specific resource.

- A simple analogy to understand security label is food labels that you see on packed food items. Some are marked with a green dot meaning that those food items are safe to be consumed by vegetarians and some are marked with a red dot making them non-vegetarian products.
- Similarly, a resource can be labelled with security classification. Some of the security labels used in the military environment are:
  - Top Secret
  - Secret
  - Confidential
  - Sensitive
  - Unclassified

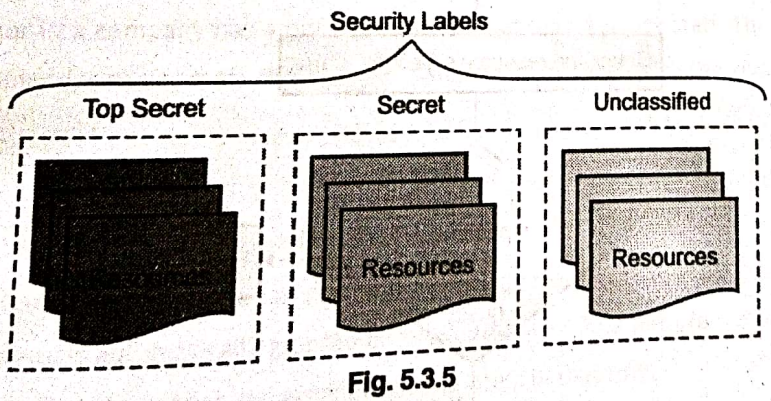


Fig. 5.3.5

Here the security labels are hierarchical in nature. This means that "Top Secret" label is a higher sensitive label than the "Confidential" sensitive label. "Unclassified" is the lowest sensitive label. If a resource is assigned "Unclassified" label, it can be accessed by anyone in the general public. Similarly, "Top Secret" is the highest sensitive label. You may use the terms security labels and sensitive labels interchangeably.

B. Clearance Level

– As security labels are assigned to the resources, clearance levels are assigned to the entities (or principals).

Definition : *A clearance level determines what resources an entity is authorised to access.*

– Similar to security labels, clearance levels used in military are:

- Top Secret
- Secret
- Confidential
- Sensitive
- Unclassified

– So, an entity with a clearance level of "Top Secret" can access all the resources that are marked with "Top Secret" security label or below. Since "Top Secret" is the highest clearance level, someone with that clearance level can access any information.

– For example, the prime minister of a country might have "Top Secret" clearance level. She will then have access to all national secret files, project plans and missions. Whereas a general government clerk might only have "Sensitive" clearance level. "Sensitive" clearance level would restrict access to resources that are marked "Sensitive" or below. Similar to security labels, clearance levels are also hierarchical in nature.

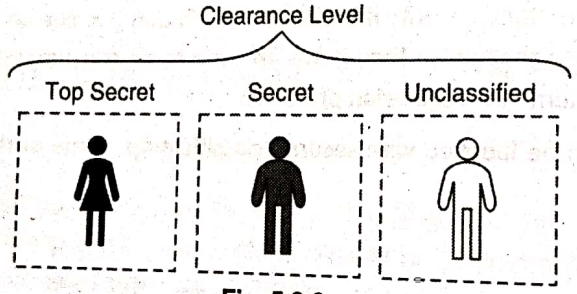


Fig. 5.3.6

C. Access Decisions in MAC

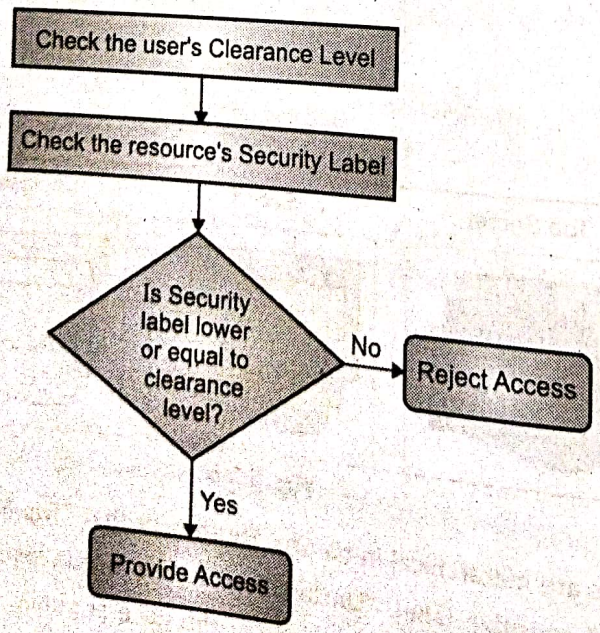


Fig. 5.3.7

- Now that you understand security label and clearance level, let us put these two concepts together to understand how MAC systems decide whether to allow access or reject it. The system first checks the clearance level of the entity requesting access to the resource. Then, the security label of the requested resource is checked.
  - If the security label of the requested resource is lower or equal to the clearance level of the requesting entity, then the access is provided.
  - If the security label of the requested resource is higher than the clearance level of the requesting entity, then the access request is rejected.
- As a rule, an entity is allowed to access resources that are marked with a security label lower or equal to its own clearance level.

Advantages of MAC

1. All resources on the system are adequately labeled.
2. Access decisions are automatically taken based on the clearance level of the requesting entity.
3. Access control policies are enforced system-wide.
4. It provides finer access policies and protects confidentiality and integrity of resources.
5. It is a scalable model for providing access control. Individual users are not required to provide access rights.

Disadvantages of MAC

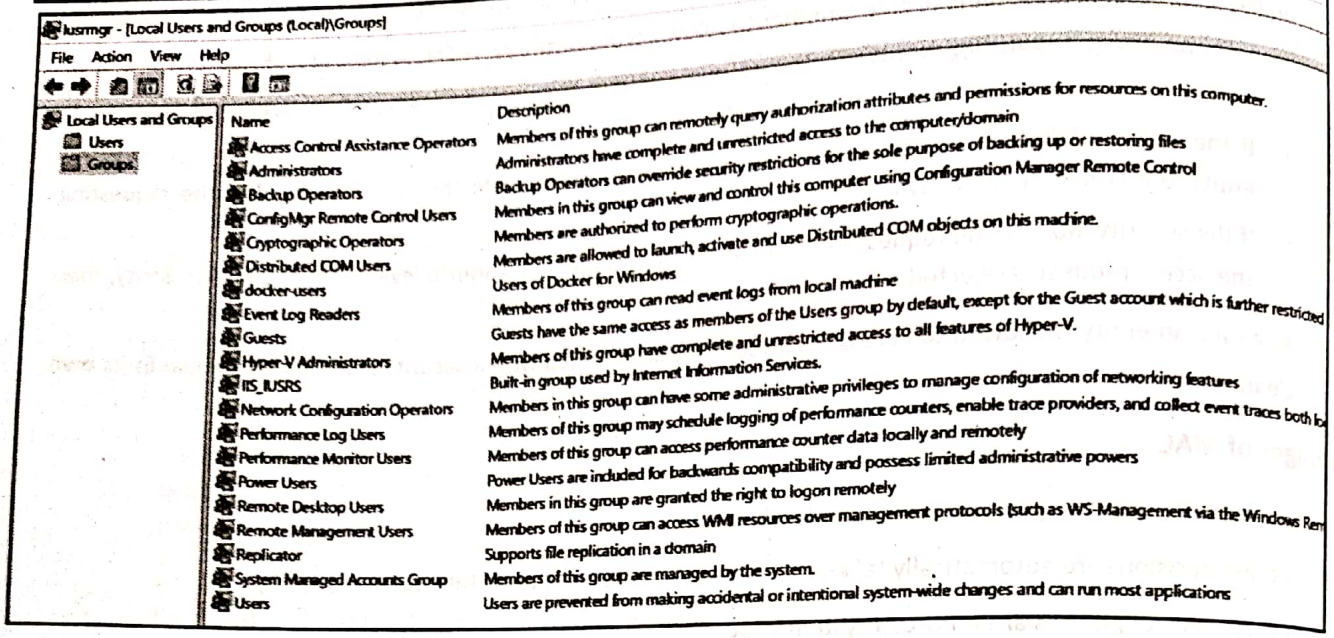
1. Difficult to configure and implement.
2. Difficult to understand compared to DAC.
3. Inflexible to provide any specialized access.

3. Role-Based Access Control (RBAC)

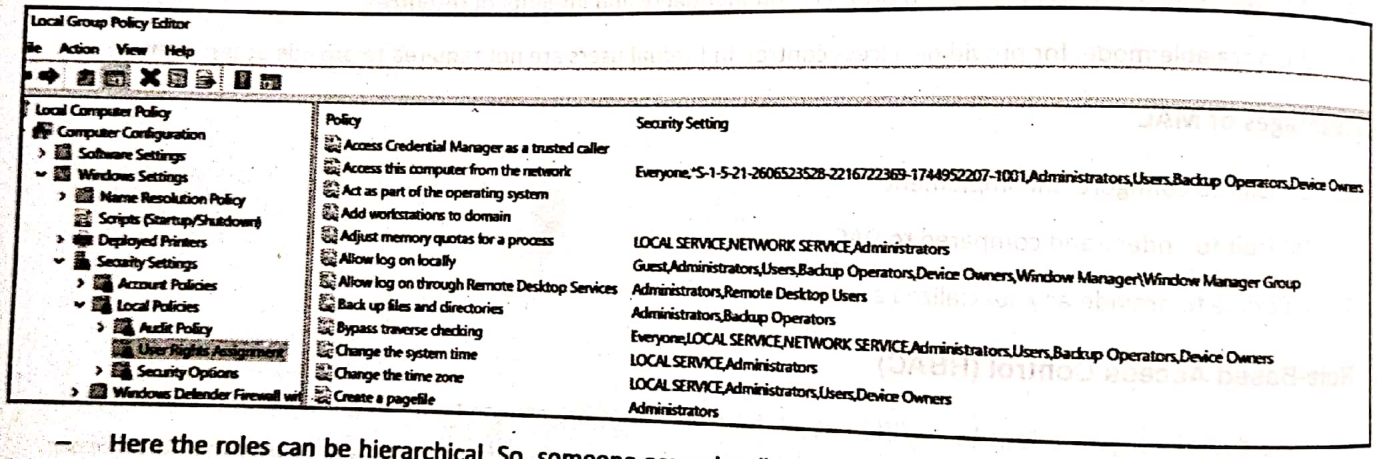
- Individuals in organisations have different job functions. They are assigned roles as per their skills and as per the organisation's requirements. For example, a college principal has a very different role to play than a professor. Similarly, a director in a company has a different role to play than a junior staff. The various duties assigned to the individuals demand various access requirements and capabilities to match the given role.

Definition : *Role-Based Access Control (RBAC) is an access control model for controlling access to resources where permitted actions on resources are identified with roles rather than with individual subject identities.*

- The objective of RBAC is to define a set of permissions for roles rather than individuals. A user, when assumes a role within the system, is automatically granted all the permissions that are pre-defined for that role. She should be able to carry out all the tasks that are assigned for the users in that role.
- For example, on a typical Windows® system, there are pre-defined groups (that work as roles).



– Each of these roles has pre-defined access rights.



– Here the roles can be hierarchical. So, someone assuming "administrator" role can pretty much do anything on the system where as "Guests" role has limited access. Users are assigned to the respective role(s). A user might be assigned more than one role on a system. Her capability to perform tasks on the system would depend on the role that she logins with.

Advantages of RBAC

1. Roles fit more naturally to the job function rather than the individual user identities.
2. Easy administration – users can come and go without needing to change the defined rights for a given role.
3. Easy to understand.

Disadvantages of RBAC

1. Role explosion – There are various job functions performed by individuals in an organisation. It is complex to define precise role requirement for each task and manage them as the organisation and business requirements evolve.
2. Not all systems support fully functional RBAC capability.

4. Attribute-Based Access Control (ABAC)

Definition : *Attribute-Based Access Control (ABAC) is an access control model in which access is decided by the policy enforced on subject and object attributes.*

Subjects (entities) and Objects (resources) are assigned various attributes. These attributes are evaluated before every requested access and the access is granted only when the attributes of both subjects and the objects match based on the pre-defined policy.

A. Attributes

Definition : *Attributes are key characteristics that can be assigned to various things to describe their state, behaviour or condition.*

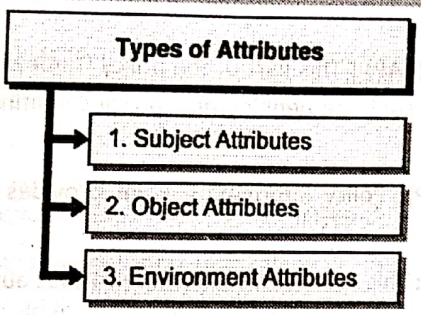


Fig. 5.3.8

There are 3 types of attributes involved in ABAC.

1. **Subject Attributes (S) :**
   These are attributes associated with a subject that defines the identity and other characteristics of that subject. Some of the attributes could be role of the subject, identity of the subject, location of the subject, age of the subject, etc.
2. **Resource Attributes (R) :**
   These attributes are associated with a resource. The resource could be a computer, file, data or anything else. The various attributes could be resource sensitivity, type of resource, location of the resource, age of the resource, etc.
3. **Environment Attributes (E) :**
   It describes the operational, technical, situational or contextual information under which access decision should be evaluated. For example, time of the day, or day of the week, etc.

B. Policy

Definition : *A policy is a set of statements or rules that specify the correct and expected behaviour of entities.*

- ABAC system uses policies to describe various attributes to evaluate access decisions. ABAC policy rules are evaluated as *IF THEN* statements.

  IF (Attributes are matched)

  THEN (allow or deny)

  ...

  ...

For example :

- Nurses in the Cardiology Department can View the Records of Heart Patients.
- Underlined words in the above policy line are various attributes that must match to allow access.
- Policy is the crucial element of any ABAC system and there are various points that manage different parts of policy components and execution.

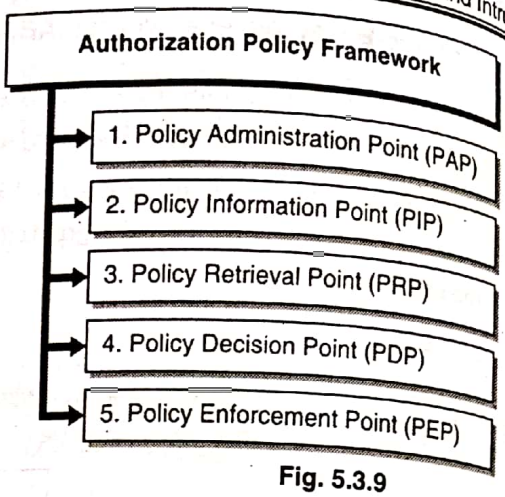


Fig. 5.3.9

1. **Policy Administration Point (PAP) :** This part of the policy framework involves general administration of various authorisation policies. These could be defining new policies, editing policies, deleting policies or archiving them for later reference.
2. **Policy Information Point (PIP) :** Policy Information Point provides attribute values for subjects, resources and environment.
3. **Policy Retrieval Point (PRP) :** This is the database where various authorisation policies are stored.
4. **Policy Decision Point (PDP) :** At this point, the various authorisation policies are evaluated. Based on the evaluation results, an authorisation decision is made either to allow or deny access.
5. **Policy Enforcement Point (PEP) :** Policy Enforcement Point receives all access requests for accessing resources from subjects. It consults PDP. If PDP allows the access, PEP lets the subject access the requested resources.

C. Access Decision in ABAC

- Let's understand how access decision is taken in ABAC.

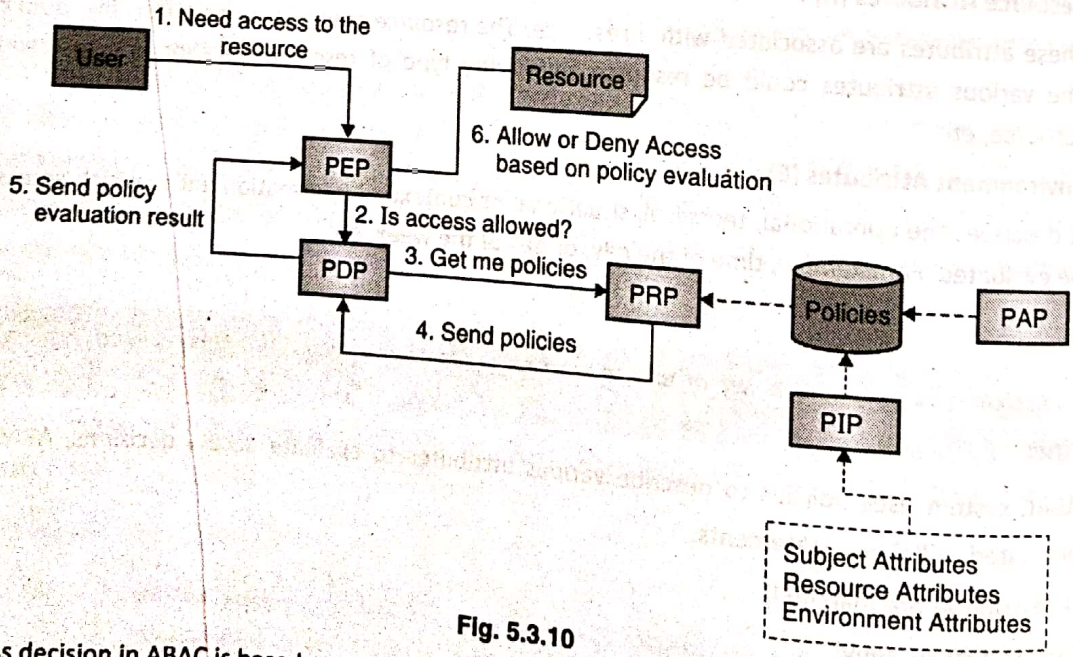


Fig. 5.3.10

- Access decision in ABAC is based on the policy definitions.
  1. The user sends the request to PEP seeking access to the desired resource.
  2. The PEP forwards the request to the PDP to evaluate if it should allow access based on the policies.

3. The PDP forwards the policy related information it requires from the PRP based on the user seeking the resource and environment conditions.
4. The PRP retrieves the various policies in the context of subject, resource and environment and sends the policies to the PDP for evaluation.
5. The PDP evaluates the policies related to the user, resource and environment conditions. It passes the evaluation result [allow access or deny access] to the PEP.
6. The PEP finally either allows access or denies it based on what it heard from PDP.

Advantages of ABAC

1. It is flexible enough to define granular conditions under which access should be granted.
2. It is scalable as the number of users and resources increases.
3. It offers a dynamic way of evaluating access decisions rather than evaluation based on identity or role.
4. It can be extended to the external users that are not part of the organisation.

Disadvantages of ABAC

1. Requires investment in infrastructure for an operational ABAC system.
2. Need to manage policies separately and periodically update them as the business environment evolves.
3. As the number of policies grows, it could be complex to understand.

5.4 Trusted Systems

Q. What is trusted system ? **SPPU - May 19**

(May 19, 4 Marks)

- Designing a secure system is a complex task. You need to follow several basic elements of security and need to model the security of the system around approved principles. These security principles are often derived and stated in mathematical formulae. One such model is the State Machine Model.

Definition : *The State Machine Model ensures that a system is in a secure state all the time. Any events that alter the state of the machine are not allowed.*

- It means that the system boots up in a secure state, carries out operations that are secure and then shutdown securely. A system is continuously in the secure state. Let us examine two such state machine models in detail.

5.4.1 Bell-LaPadula (BLP) Model

- The Bell-LaPadula (BLP) Model was developed in 1970s for the US military to avoid disclosure of sensitive information. It addresses confidentiality aspects of access control and ensures that the sensitive information is not leaked. Its application has been mathematically proven to ensure confidentiality.

Definition : *The Bell-LaPadula Model is a state machine model that enforces confidentiality aspects of access control.*

- Like Mandatory Access Control, BLP model uses security levels for classifying subjects and objects.

Definition : *A system that uses security levels for classifying subjects and objects is called a multilevel security system.*

- The BLP model has 4 core components.
  1. **Subjects :** Entities that desire access to the resources
  2. **Objects :** The desired resources [file, information, server, etc.]
  3. **Actions :** Read, Write
  4. **Security Levels :** Sensitivity labels [for example, top secret, secret, confidential, sensitive, unclassified]
- The BLP model provides four rules for access control.

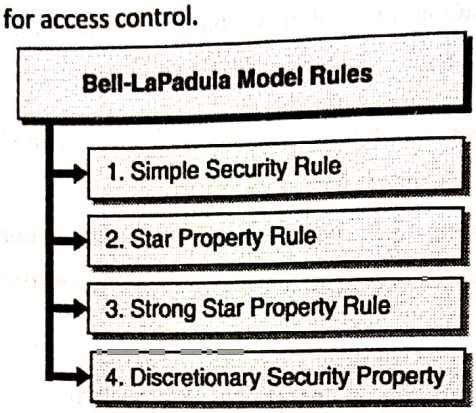


Fig. 5.4.1

1. Simple Security Rule

Definition : *The Simple Security Rule in the Bell-LaPadula Model states that a subject cannot read an object at a higher security level than itself.*

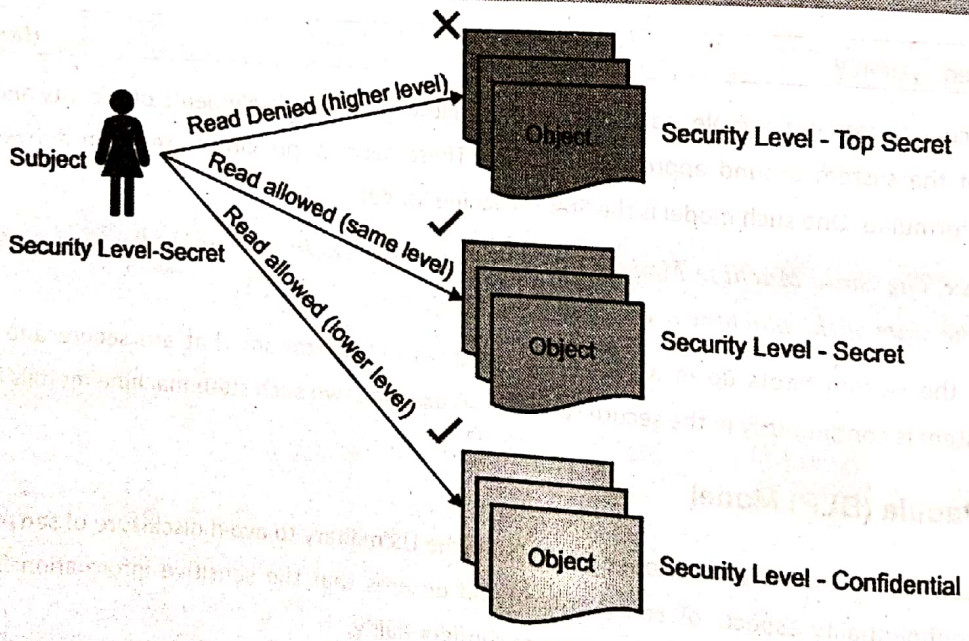


Fig. 5.4.2

- This rule is also called "No-Read Up" or simple security property (ss-property). What this means is that any subject must have a security level greater than or equal to the security level of the object it is trying to access. A subject cannot read an object at a higher security level than itself.

- For example : take a practical scenario. Does everyone in a country can get access to the military plans and missions? Why not? The security level of military information is so high that general citizens do not have adequate clearance level to get access to it.
- Take another example of your house. Can everyone access your locker? Why not? The trust level (or the security level) of everyone is not high enough to access the locker, isn't it? That's precisely what the simple security rule tells you. Do not allow an entity at a lower trust level to access higher trust level objects (or resources).

2. Star Property Rule

Definition : The Star Property Rule in the Bell-LaPadula Model states that a subject cannot write to an object at a lower security level than itself.

- This rule is also called "No-Write Down". What this means is that no subject can write to objects that are at a lower security level than itself. However, writing to objects at a higher security level is allowed.
- Let's take an example. Did you hear about the demonetization of 8 November 2016 before the PM announced it? Perhaps, no. Who else do you think would know about such a move? I hear you saying that very highly placed and trusted sources. Exactly. So, subjects that have access to highly trusted information are not supposed to write it down (or share) with someone who is not at an equally trusted level.
- That's precisely what the Star Property Rule in the BLP model tells you. Do not share (or write) higher security level information with an object at a lower security level. If you do so, subjects at a lower security level can read such objects and can know about the information you wrote that was not supposed to be shared. Remember that the BLP model focuses on confidentiality. Such, lower level sharing could mean disclosure and hence the Star Property Rule of the BLP model denies it.

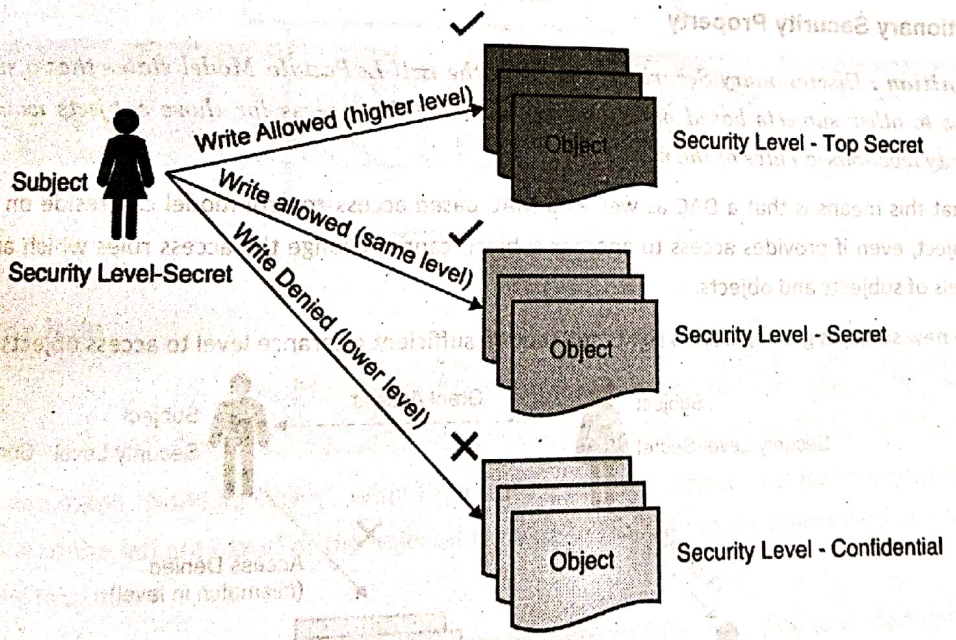


Fig. 5.4.3

3. Strong Star Property Rule

Definition : *The Strong Star Property Rule in the Bell-LaPadula Model states that a subject can only read and write to an object at the same security level as itself.*

- You can consider this as a combination of first two rules – Simple Security and Star Property. Simple Security Rule deals with Read operations whereas the Star Property Rule deals with Write operations.
- So, if a subject has to do both – read and write – combining those two rules – gives permissions to do so only if the subject and the object are at the same security level.

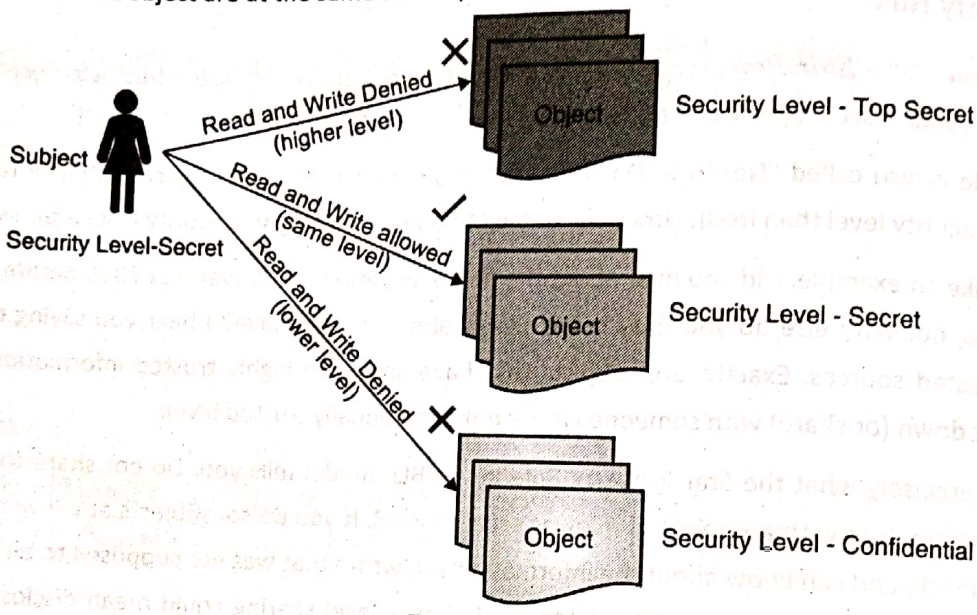


Fig. 5.4.4

4. Discretionary Security Property

Definition : *Discretionary Security Property in the Bell-LaPadula Model states that a subject can provide access to other subjects based on its own judgement. The access for those subjects would still follow the security level-based rules of the model.*

- What this means is that a DAC as well as a MAC based access control model can reside on the same system. A subject, even if provides access to another subject, cannot change the access rules which are based on security levels of subjects and objects.
- The new subjects gaining access rights still require sufficient clearance level to access objects.

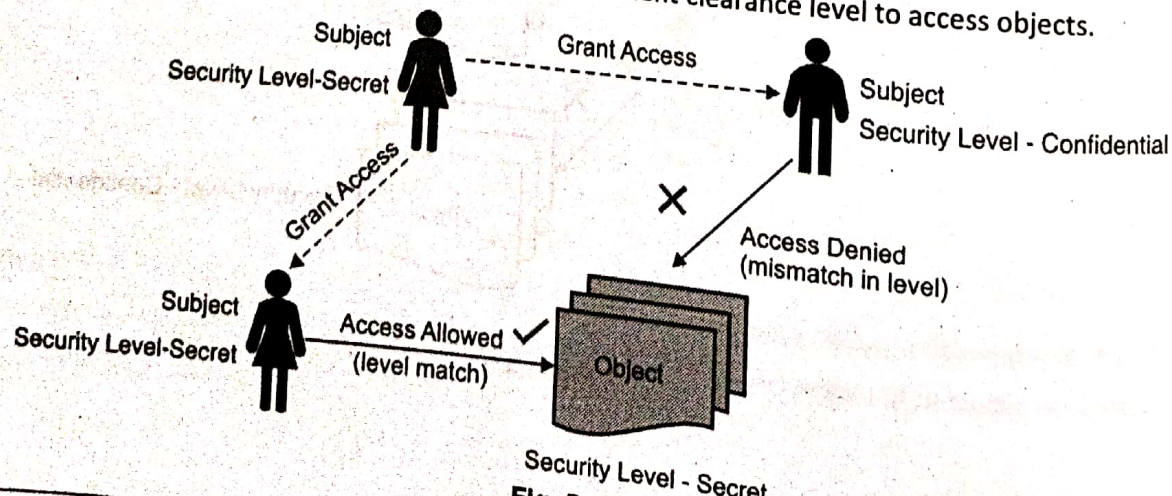


Fig. 5.4.5

Summary

Sr. No.	Rule Name	Purpose
1.	Simple Security	Secure Read Operations (no read up)
2.	Star Security Property	Secure Write Operations (no write down)
3.	Strong Star Property	Secure Read and Write Operations (read and write at same level only)
4.	Discretionary Property	Access Grant Operations (first 3 rules apply even if access is granted)

5.4.2 Biba Model

The Biba Model was developed in 1975 to avoid alteration of information. It addresses integrity aspects of access control and ensures that the information (data) is not altered by unauthorised entities. Its application has been mathematically proven to ensure integrity.

Definition : The Biba Model is a state machine model that enforces integrity aspects of access control.

Unlike the BLP model, the Biba model uses integrity levels for classifying subjects and objects. Integrity levels can be thought to follow similar hierarchy as security levels. It is just that it is called integrity level instead of security level when referring to the Biba model.

The Biba model provides three rules for access control.

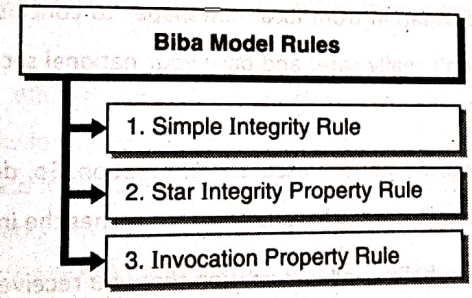


Fig. 5.4.6

1. Simple Integrity Rule

Definition : The Simple Integrity Rule in the Biba Model states that a subject cannot read an object at a lower integrity level than itself.

- This rule is also called "No-Read Down". What this means is that any subject must have an integrity level greater than or equal to the integrity level of the object it is trying to read. A subject cannot read an object at a lower integrity level than itself.
- Confused? Do not compare the rule with the BLP model. Here you are not trying to protect confidentiality but ensuring integrity of information.

- Let's take an example. Suppose that you are the Prime Minister of a country. Should you attack on a country by just reading some article in a local newspaper that is speculating about the plans of the enemy country to attack your country? Probably not. You will check the information from your security agencies.
- The information coming from national security agencies would be more accurate and factual rather than speculative.

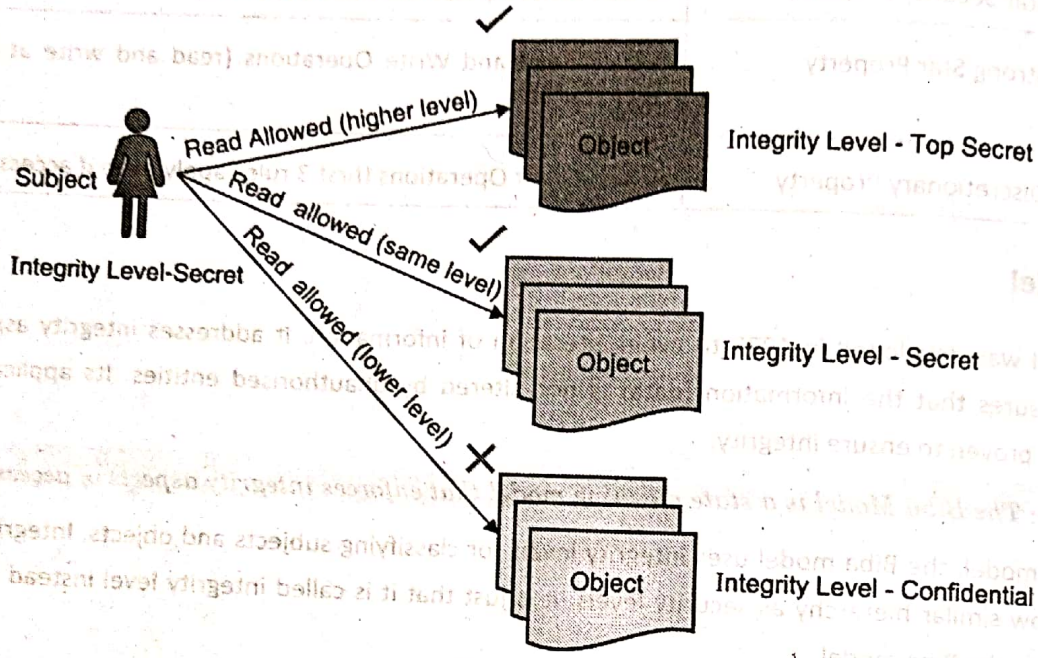


Fig. 5.4.7

- You won't read and consume information from local newspaper to conclude on your decisions. So, you, being Mr. PM, won't read (or rather won't really care) and base your national security decisions on articles at a lower integrity (assurance) level.
- You at your level really need to know (and get) accurate information. So, don't read articles that are at a lower integrity level. Maintain the source of information such that it matches the integrity level you are at.
- Let's take another example. Do you believe all information that you receive on WhatsApp or other social media platform? Haven't you encountered fake messages or rumours anytime? Don't you feel that all information on social media platform is not accurate and there is a need to validate the truth (integrity) of the information before believing it? Don't you check for the validity of the information from sources which are more trusted such as news channels and newspapers? That's precisely what Biba model tells. The information you consume should match your integrity level.

2. Star Integrity Property Rule

Definition : *The Star Integrity Property Rule in the Biba Model states that a subject cannot write to an object at a higher integrity level than itself.*

- This rule is also called "No-Write Up". What this means is that a subject is not allowed to write information to the objects that are at the higher integrity level. What would happen if such a write is allowed? The lower integrity subject can corrupt the higher integrity information.

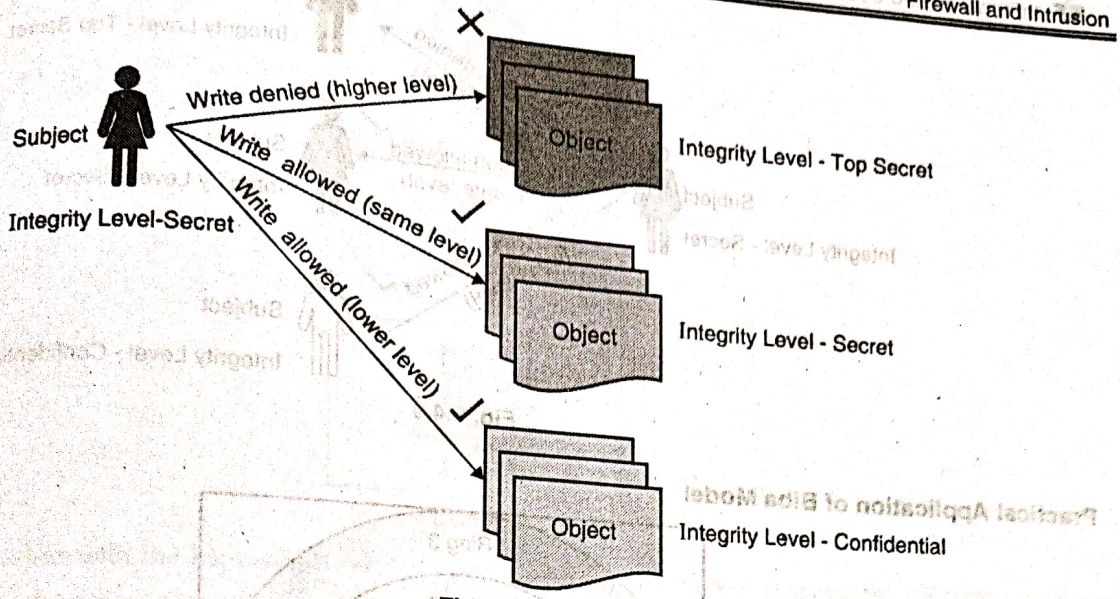


Fig. 5.4.8

- Again, don't be confused. You are not protecting confidentiality here but integrity. Let's take an example. Would a national level newspaper allow printing of national level news from anyone? Perhaps no. it would doubly check even from its most trusted journalists that the news is accurate to the best of their knowledge. So, basically, when someone is writing to a much valued (trusted) object, that source (or subject) who is writing it must be trusted enough to allow such a write. You do not allow anyone and everyone to write information on such trusted objects.
- Let's take another example. Who should be feeding exam results into the results portal? Would an institute allow everyone to enter the marks in the result portal? Why not? The reason again being that the institute does not want to corrupt the legitimate marks information. Someone who is at a lower integrity level might pass someone who is actually a fail. The marks can be increased or decreased as well during entering them in the results portal. So, the results portal, that is trusted by thousands of candidates, would not and should not display incorrect results. The integrity of results is much more important than the confidentiality of the results. That's what Biba's Star Integrity Property Rule tells you.

2. Invocation Property Rule

Definition : *The Invocation Property Rule in the Biba Model states that a subject cannot communicate with another subject at a higher integrity level than itself.*

- Let's take an example. Will an army chief work on instructions from a commander? Will a principal work on instructions from a peon? Will You work on instructions from a stranger without any authority? I think you get it now.
- The Invocation Property Rule allows a subject to communicate with another subject only if the subjects are at the same or lower integrity levels. A lower integrity level subject cannot communicate (or invoke or call) a higher integrity level subject.

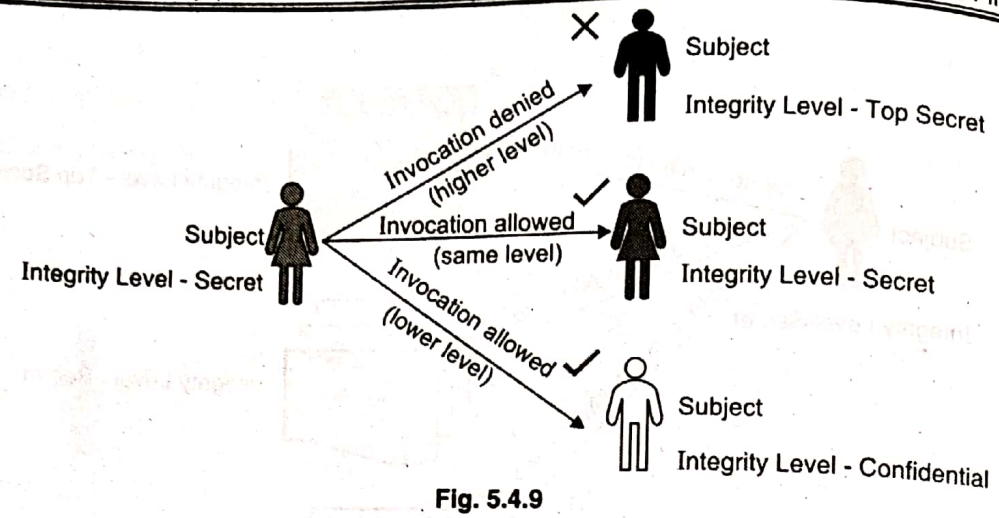


Fig. 5.4.9

Practical Application of Biba Model

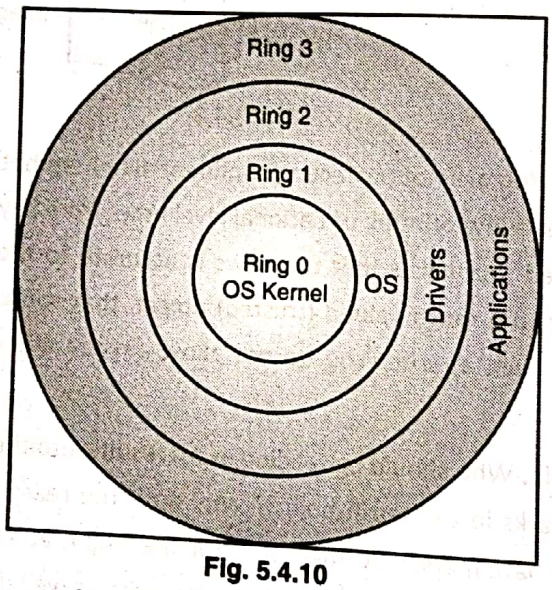


Fig. 5.4.10

- So, where is Biba model seen these days? In the CPU and Operating System architecture.
- The CPU typically has a protection ring architecture whereby the most trusted components of the system have access to higher privileges (permissions), and the untrusted components are executed in the outer ring with lesser privileges.
- The OS and CPU developers do not know in advance that what are application developers going to do. Hence, the protected ring mechanism ensures that the higher privilege commands and instruction sets are not available to the applications. Applications if need to execute such commands then they must make a system call to the operation system which can execute the commands on the behalf of such applications.
- Similarly, Operating Systems (OS) put strict controls in terms of what a user or application can do. OS is responsible for managing all interactions with the hardware on a computer and thus is directly responsible for managing its logical security. It also employs the layered architecture to protect the information systems.
- The outer most layer is User Space. In this space, the user installs applications and interacts with other functions of the operating system such as playing a music or storing a file. Space below user space is called the kernel space. This is operating system's private area. User or applications cannot directly interact with the OS in this area.

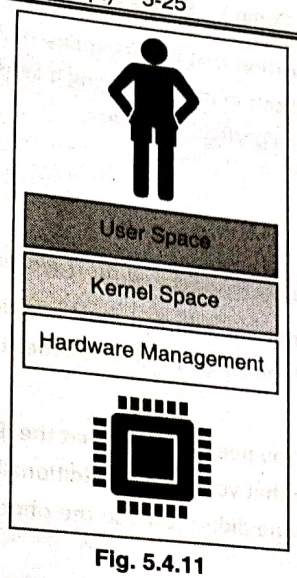


Fig. 5.4.11

- Also, interaction with the hardware is restricted via the OS only. In the kernel space, the OS performs various functions such as Input Output Management, Memory Management, Disk Management, CPU process handling and other general hardware functions.
- So, to maintain the overall system integrity the entities are controlled according to the Biba model. Entities are only allowed to read and write to what their integrity level permits. So, for example, a virus program is restricted to write to the kernel memory.

Summary

Sr. No.	Rule Name	Purpose
1.	Simple Integrity	Secure Read Operations (no read down)
2.	Star Integrity Property	Secure Write Operations (no write up)
3.	Invocation Property	Secure Communications (no interaction with higher level subject)

5 Authentication Methods

Before you jumpstart on authentication methods, let's refresh a few terms and concepts around it.

5.1 Introductory Concepts

Identification

- Identification (in short ID) is defined as,

Definition : A way to claim an entity's presence with respect to the process being carried out.

- This means that during a process, your presence (or your consent) is ascertained (or established). For example, when you try to login to your Facebook account, you provide your Email or Phone number to establish your presence during the login process.

- There are several other forms of identification that we use today such as Aadhar Card, PAN Card, Voter ID, Debit Card, Admit card, etc. All of these identification methods bring a sense of credibility that you are present, or you give your consent (approval) to complete a particular process.

2. Authentication

- Authentication is defined as,

Definition : *A way to ensure that the entity is indeed what it claims to be.*

- This means that providing just the ID is not enough. You must additionally prove that the ID belongs to you. For example, even if I know your Facebook email address or phone, I cannot login as you until I also know the password.
- Thus, knowing just the ID is not enough. You need to prove that the ID belongs to you and that is what is precisely called authentication. It is for this reason that you need to additionally sign when you submit Aadhar card or PAN card as an ID proof to ensure that someone didn't just use the photocopy of those IDs without your permission (or consent).
- Some of the ways to authenticate an ID are passwords, biometric (like your Aadhar fingerprints or phone sensor), PIN (like for Debit Card), or OTP (SMS that you get to confirm a transaction).

3. Authorisation

- Authorisation is defined as,

Definition : *A way to determine what resource an entity can access.*

- Once you have provided your ID and have been successfully authenticated, the next step is authorisation where the system determines if you have the permission to access the desired object. For example, even if you have a valid voter ID card but if your name is not on the electoral list at a particular area booth, you won't be allowed to vote. Having authenticated ID is one thing and getting access to the resource is another.
- Just because you have an authenticated ID, does not mean that you have automatically access to the resources. So, authenticated ID is a must for authorisation but that does not always guarantee that you would be allowed access.

4. Accountability

- Accountability is defined as,

Definition : *A way to record your actions.*

- Suppose, you used a system to take print outs. That system logs this action (pretty much like you record attendance in lab or classroom) to build a trace (evidence or proof) that you used the printer.
- If you were not supposed to use the printer, the evidence can be used to find you accountable for using it without permissions and could result in particular consequences.
- Accountability is a key determinant of how securely a system is operating. The logs generated are continuously monitored and necessary alarms are raised if any entry is found to be suspicious.
- Let's summaries the 4 access control steps with the help of Fig. 5.5.1.

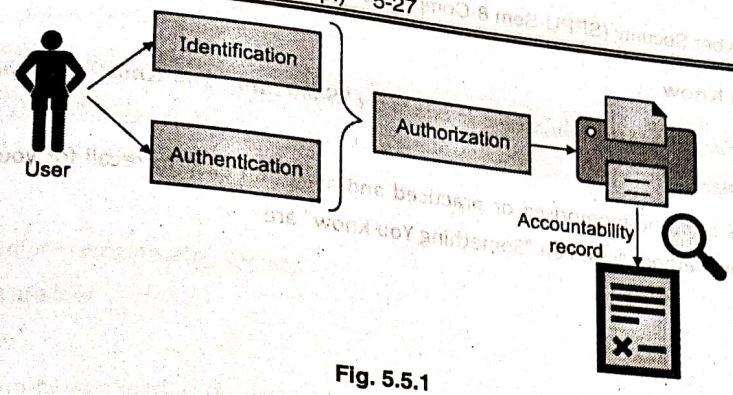


Fig. 5.5.1

Non-repudiation

- Non-repudiation is defined as,

Definition : A way to prove your actions.

- It is used in conjunction with authentication and accountability. Non-repudiation provides an assurance that someone cannot deny their actions later on. For example, if I sent you an email, I cannot later deny that I did not.
- To send an email, I must have used my email ID and password and then sent it over to you over a secure network where no one could change the email body. If you can establish all of these facts truthfully, you have proven that I sent that email and thus established non-repudiation.

5.5.2 Types of Authentication Methods

- Overtime, various authentication methods have evolved to address
  - the ease of authentication
  - make it hard to break authentication
  - make authentication techniques suitable for various devices
- At a high level, authentication methods are categorised as shown in Fig. 5.5.2.

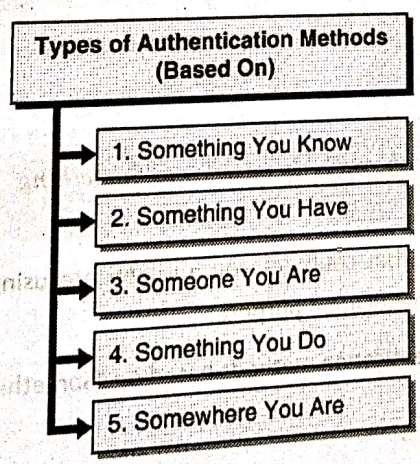


Fig. 5.5.2

1. Something You Know

Definition : *The authentication methods based on "Something You Know" rely upon your secret knowledge about something.*

- These secrets could be memorised or practiced and are often easy to recall for you. Some of the examples of authentication methods based on "Something You Know" are
- Passwords
- PIN
- Passphrases
- Secret questions
  - Mother's maiden name
  - First pet name
  - First car purchase year
  - First school name
  - City in which you were born
  - Or other secret questions whose answers would be likely known just to you
- Lock key combination

Advantages of "Something You Know" type authentication

- Easy to implement by developers in the product (OS, Applications or Websites)
- Easy to recall for the user
- Easy to authenticate for the user
- Easy to change for the user
- Chances of errors are low
- Can be used by non-person entities (hardware devices, applications, OS, programs, processes, etc.)

Disadvantages with "Something You Know" type authentication

- Easy to crack (or break)

2. Something You Have

Definition : *The authentication methods based on "Something You Have" rely upon your possession of something.*

- You could possess something that could let you authenticate using it. You often don't require to remember anything while using it for authentication purpose.
- Some of the examples of authentication methods based on "Something You Have" are
  - Physical keys
  - Badge
  - Swipe Card (for example Debit or Credit Card)
  - Digital Certificates

- Security Keys (for example Private Key)
- OTP (that you get via SMS)
- Tokens

Advantages of "Something You Have" type authentication

- Easy to use
- Does not often require remembering secrets
- Chances of errors are low
- Not easy to crack
- Can be used by non-person entities (hardware devices, applications, OS, programs, processes, etc.)

Disadvantages of "Something You Have" type authentication

- Difficult to change
- Possibility of loss or theft
- Requires distribution methods (provisioning) to reach to the user securely

3. Someone You Are

Definition : *The authentication methods based on "Someone You Are" rely upon your physical characteristics.*

- Your body has several physical characteristics that can be used to uniquely identify and authenticate you. These characteristics do not much change over time (as you age) and can serve authentication purpose for near lifetime.
- These characteristics are called Static Biometrics.

Definition : *Static Biometrics are physical characteristics that can be used for authentication.*

- Generally speaking,

Definition : *The measurement and analysis of unique physical or behavioural characteristics is called biometrics.*

- However, there can be scenarios where re-provisioning (re-calibrating) your biometric details might be required. For example, if you use a particular finger for fingerprint and your that particular finger is damaged permanently, you might have to choose another finger or another type of biometrics for authentication.
- Some of the examples of authentication methods based on "Someone You Are" are
  - Fingerprint
  - Palm Scan
  - Hand Geometry
  - Retina Scan
  - Iris Scan
  - Facial recognition

Advantages of "Someone You Are" type authentication

- Easy to use
- Does not often require remembering secrets
- Difficult to crack

Disadvantages of "Someone You Are" type authentication

- Difficult to implement correct
- Chances of errors are high (recall you trying fingerprints several times at Aadhar enrolment centre?)
- Difficult to change (requires physical presence for re-provisioning)
- Cannot be used by non-person entities (hardware devices, applications, OS, programs, processes, etc.)

4. Something You Do

Definition : *The authentication methods based on "Something You Do" rely upon your way of performing a given task.*

- This authentication method also collects biometric patterns but while performing a given task. Unlike static biometric patterns, these are dynamic biometric patterns which are used for authentication purpose.

Definition: *Dynamic Biometrics are job performing characteristics that can be used for authentication while an individual is performing a given task.*

- Some of the examples of authentication methods based on "Something You Do" are
  - Voice print (or pattern)
  - Keystroke Dynamics (how hard you press the keys and how fast)
  - Handwriting characteristics (remember old movies where handwriting matches were done?)

Advantages of "Something You Do" type authentication

- Easy to use
- Does not often require remembering secrets

Disadvantages of "Someone You Are" type authentication

- Difficult to implement correct
- Not too difficult to crack (consider replaying a recorded voice)
- Chances of errors are high (have you tried voice to text yet?)
- Difficult to change (requires physical presence for re-provisioning)
- Cannot be used by non-person entities (hardware devices, applications, OS, programs, processes, etc.)

5. Somewhere You Are

Definition : *The authentication methods based on "Somewhere You Are" rely upon your physical location.*

- Increasingly, the devices and systems have location awareness. Your mobile phone or laptops (via network connectivity) precisely know where you are located at a particular moment. "Somewhere You Are" uses the location information for authentication.

- For example, Google Smart Lock for Android allows you to set Trusted Locations. Say, if you are at home, your phone may not require you to unlock it (via PIN, pattern, password or fingerprint) before use.

https://support.google.com/accounts/answer/6160273?hl=en

Set your Android device to automatically unlock

You can keep your Android phone or tablet unlocked in some situations, like when your phone is in your pocket or you're near home. When you use Smart Lock, you won't need to unlock with your PIN, pattern, or password. The features you can use depend on your device.

If you want to change your screen lock, learn more about screen lock settings.

Note: Some of these steps work only on Android 9 and up. Learn how to check your Android version.

Turn on automatic unlock

1. Make sure you have a screen lock. Learn how to set a screen lock.
2. Open your device's Settings app ⚙.
3. Tap **Security & location** > **Smart Lock**.
4. Enter your PIN, pattern, or password.
5. Pick an option and follow the on-screen steps.

After setup, when you turn on your screen, you'll see a pulsing circle at the bottom around the Lock 🔒.

Important: When you don't use your device for 4 hours, and after it restarts, you'll need to unlock it.

Turn off automatic unlock

1. Open your device's Settings app ⚙.
2. Tap **Security & location** > **Smart Lock**.
3. Enter your PIN, pattern, or password.
4. Turn off **On-body detection** and remove all trusted devices, trusted places, trusted faces, and Voice Match voices.
5. Optional: If you want to turn off your screen lock, learn how to change your screen lock.

- "Somewhere You Are" is also quite widely used in corporate IT. In many environments, if you are on an office network (LAN or Wi-Fi), you can login using only a password, but if you are out of the office you must use VPN or an additional mechanism for authentication.

Advantages of "Somewhere You Are" type authentication

- Easy to use
- Does not often require remembering secrets
- Can be used for authenticating non-person entities (mobile, laptops or other location aware devices)

Disadvantages of "Somewhere You Are" type authentication

- Not too difficult to crack (consider theft and someone carrying the device to a trusted location)
- Cannot be used for authenticating individuals
- Requires network connectivity for location awareness

5.5.3 Comparison between the AuthenticationTypes

Sr. No.	Type of authentication	Ease of use	Ease of Change	Ease of implementation	Error rate	Support non-person entities
1.	Something You Know	High	High	High	Low	Yes
2.	Something You Have	High	Low	Medium	Low	Yes
3.	Someone You Are	High	Low	Low	High	No
4.	Something You Do	Medium	Low	Low	High	No
5.	Somewhere You are	High	High	Medium	Medium	Yes

5.5.4 Factors of Authentication

Each type of authentication that you learnt in this section is called a factor of authentication. Based on the number of factors you choose for effectively carrying out and completing the authentication process, you have three types of authentication scenarios.

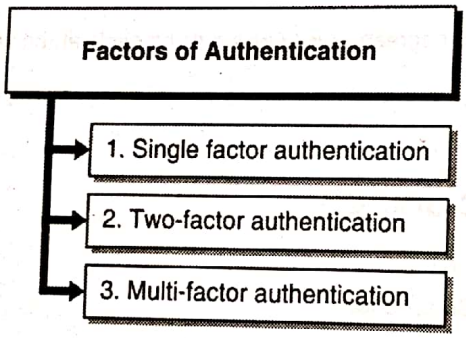


Fig. 5.5.3

1. Single factor Authentication

✍ ***Definition :*** *Single factor authentication requires only ONE of the types of authentication for successfully carrying out the authentication process.*

For example, you could just use password or token. This is most widely used. It is often treated as a weak form of authentication.

2. Two-factor Authentication

✍ ***Definition :*** *Two-factor authentication requires you to use any TWO types of authentication, one after another, for successfully carrying out the authentication process.*

- It is considered as a strong form of authentication. For example, for your online transactions, you are first required to give the account password and then you receive an OTP. You are required to put the correct OTP for successfully authenticating your account details and carrying out and completing your transaction.
- Another example could be ATM. You must possess you Debit Card (Something You Have) and put in the right PIN (Something You Know) for withdrawing money.

Multi-factor Authentication

Definition : Multi-factor authentication requires you to use MORE than two types of authentication.

It is often used in a high security environment. For example, you may be first required to give your fingerprint, after which you can access an application where you are required to provide username and password. For carrying out a transaction on that application, you might require an OTP.

Note : If two authentication techniques from the same type of authentication are used, it is not considered two-factor authentication. It would be considered a single factor authentication. For example, you cannot consider consecutive requirement of two passwords or a password and a PIN to be two-factor authentication. Two-factor authentication essentially requires two different ways to authenticate.

5.5.5 Password Based Authentication

Definition : A password is a protected sequence of characters that is used to authenticate an entity and provide access to a protected system.

Passwords are the most widely used form of authentication. You can use it on computers, mobile devices, web portals, applications and nearly everything else that supports any form of user interface and requires authentication before use.

Passwords fall into "Something You Know" type of authentication that you learnt earlier. The use of passwords requires effective password management that is strong enough to keep the passwords protected.

Choosing a Password

Passwords, even though are the easiest to use amongst other types of authentication techniques, are also weakest form of authentication. Why? Because, users usually choose weak passwords which are easily breakable or guessable.

Passwords can either be user chosen or be automatically generated by a trusted system.

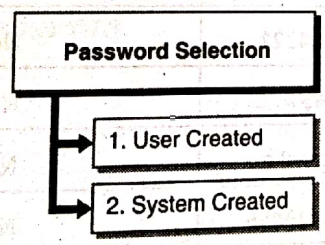


Fig. 5.5.4

User Created

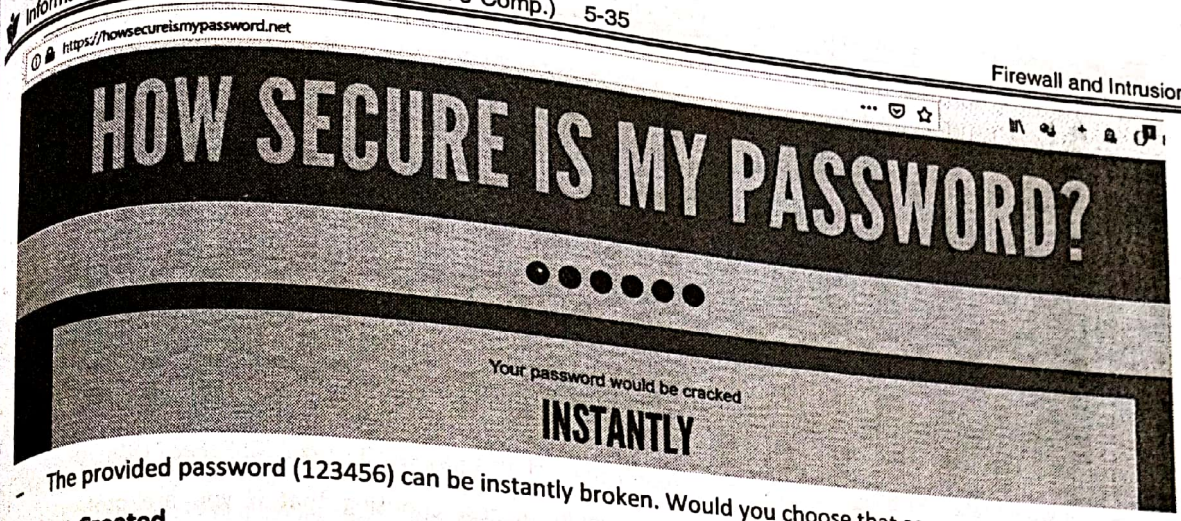
A user can typically choose her own password when signing up on a system. The problem here is that the users typically choose something that is extremely easy to guess.

Some of the common and worst password choices of 2018 are as listed in Table 5.5.1.

Table 5.5.1 : Source : SplashData's Top 100 Worst Passwords of 2018

RANK	PASSWORD	CHANGE FROM 2017
1	123456	Unchanged
2	password	Unchanged
3	123456789	Up 3
4	12345678	Down 1
5	12345	Unchanged
6	111111	New
7	1234567	Up 1
8	sunshine	New
9	qwerty	Down 5
10	iloveyou	Unchanged
11	princess	New
12	admin	Down 1
13	welcome	Down 1
14	666666	New
15	abc123	Unchanged
16	football	Down 7
17	123123	Unchanged
18	monkey	Down 5
19	654321	New
20	!@#$%^&*	New
21	charlie	New
22	aa123456	New
23	donald	New
24	password1	New
25	qwerty123	New

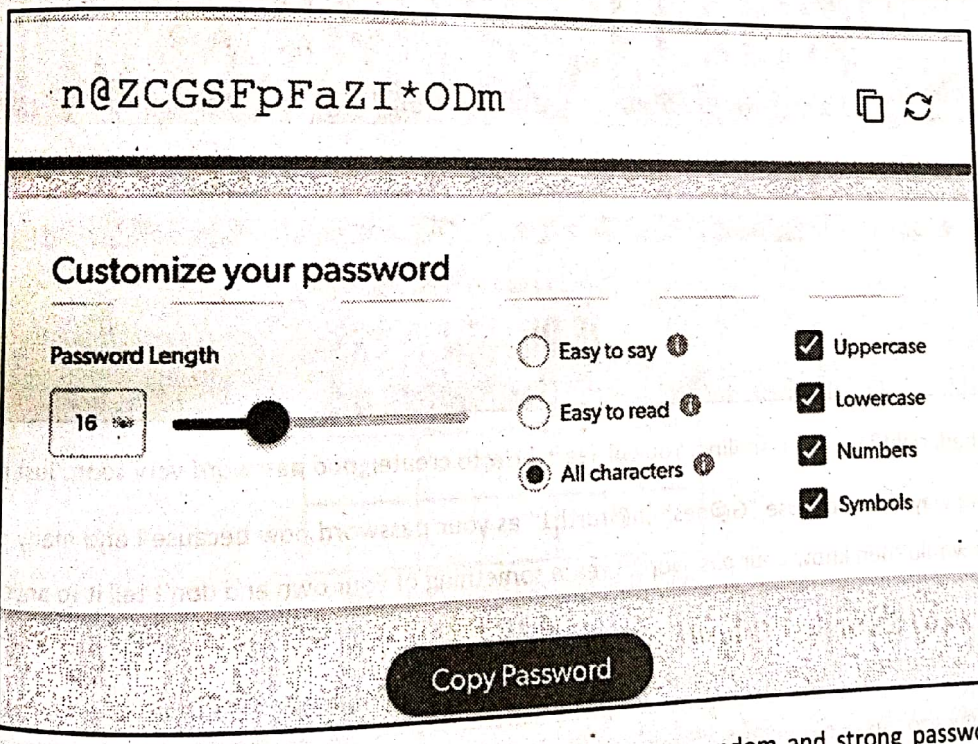
- Is your password on the list?
- Let's pick the first password from this list (123456) and find out how much time it would take a computer to break this password. https://howsecureismypassword.net/ provides a good estimate of password cracking times with respect to the provided password.



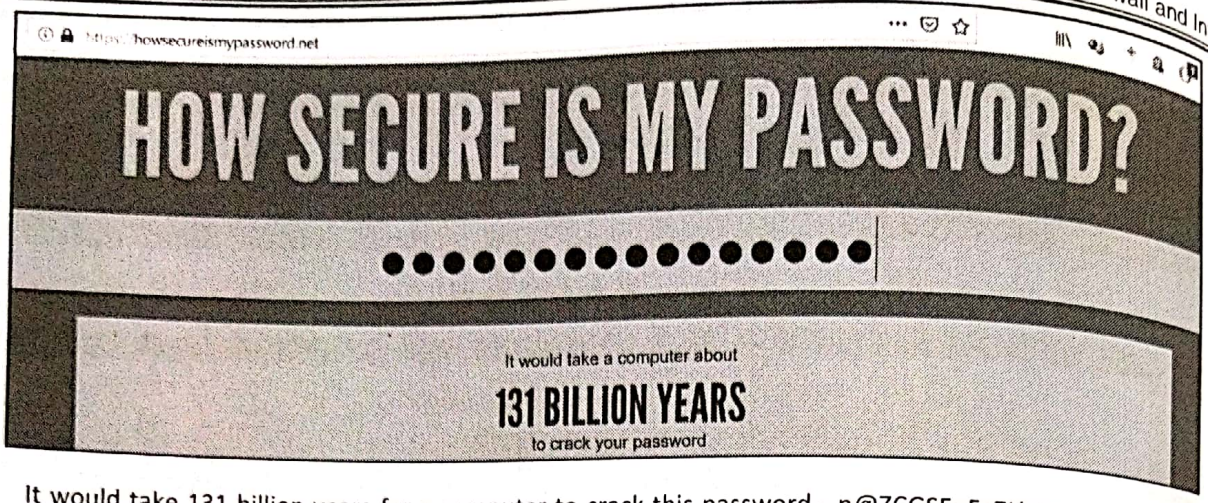
- The provided password (123456) can be instantly broken. Would you choose that as your password still?

2. System Created

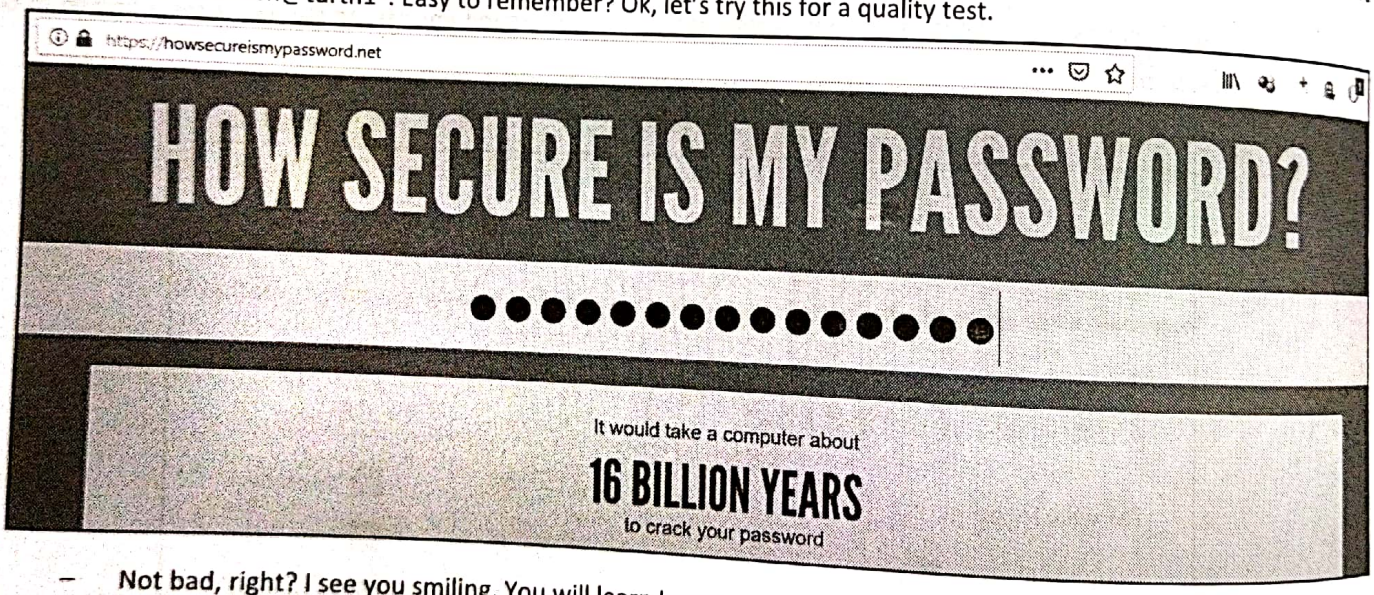
- Quite a few systems today offer passwords that are
  - o Randomly generated
  - o Difficult to crack and
  - o Easy to remember
- Such systems ensure that the important quality parameters for a good and strong password are matched in the suggested passwords.
- Here is a snapshot of LastPass password generator tool.



- It allows you to choose length, mix of characters and creates a random and strong password. Let's find the strength of the password it generated (n@ZCGSFpFaZI*ODm).



- It would take 131 billion years for a computer to crack this password - n@ZCGSFpFaZI*ODm. Wow, now that is something called as a strong password! I would suggest choosing that if you are protecting something worthwhile!
- But, do you say that the password (n@ZCGSFpFaZI*ODm) is hard to remember? Ok, I hear you.
- Let me give you a tip. How about choosing an Indian festival as a password but slightly altering the characters? Let's take "Ganesh Chaturthi". I will replace "a" with "@" and "i" with "1". So, "Ganesh Chaturthi" would be "G@neshCh@turth1". Easy to remember? Ok, let's try this for a quality test.



- Not bad, right? I see you smiling. You will learn how to create good password very soon. Just read on!
- By the way, don't choose "G@neshCh@turth1" as your password now because I and many other readers of this book would then know your password. Create something of your own and don't tell it to anyone.

5.5.6 Password Selection Criteria (Quality Guidelines)

Q. Explain any two password management practices. **SPPU – May 19**

(May 19, 4 Marks)

Here are some general guidelines when choosing passwords.

1. Choose at least 12 characters in your password. The lengthier the password, the harder it is to crack it.

2. Use mix of characters.
   a. At least one digit.
   b. At least one uppercase letter.
   c. At least one special character such as (@,#, $, %, ^, &, *, (,), etc.).
3. Do not use the same password for all your accounts.
4. Do not use easy to guess passwords. Do not use your name, date of birth, school name etc. as your password.
5. Check your password quality before using it.
6. Do not use words found in the dictionaries. For example, avoid creating a password such as "Apple".

General Password Usage Guidelines

Here are some general guidelines to follow when working with passwords.

1. Do not send your password in cleartext. Avoid entering them on sites that do not use https.
2. Do not share your password with anyone in your family and friends.
3. Change your passwords periodically to avoid overuse.
4. Do not tell your password to anyone over phone or email howsoever the conversation may sound legitimate.
5. Wherever possible, use two-factor authentication with password.
6. Do not write down your password!

5.5.7 Storing Passwords on System

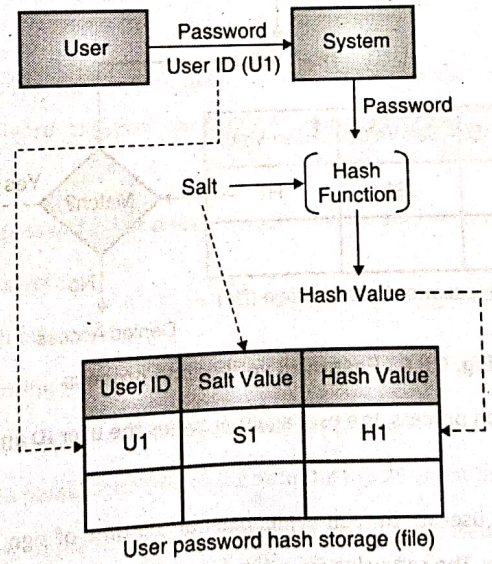


Fig. 5.5.5

- When you enter your password on a Windows® or a Linux machine, how does the system know that you entered the right password? Does it compare your provided password, character by character, to the one previously stored on the system that you chose during sign up? No, it does not store your password in cleartext at anywhere on the system.
- The passwords are adequately protected using hashes.

A. Sign up Process

- During the sign-up process, the user creates an ID and chooses a password that she would like to use when authenticating to the system.
- The system chooses a random salt value (nonce or any other random value) and passes that along with the user provided password to a hashing function. The hashing function computes the hash value. Then, the hash value and the corresponding salt value with the user ID is saved on the system in the form of a file. The system does not store password in cleartext at anywhere on the system.
- The purpose of the salt value is to bring randomness to the hashing process. So, if any two users provide the same password, even then, the entries in the password hash table file would not be the same. The random salt value also serves a purpose of defending the passwords from attack as you would see later in the section.

B. Password Verification Process

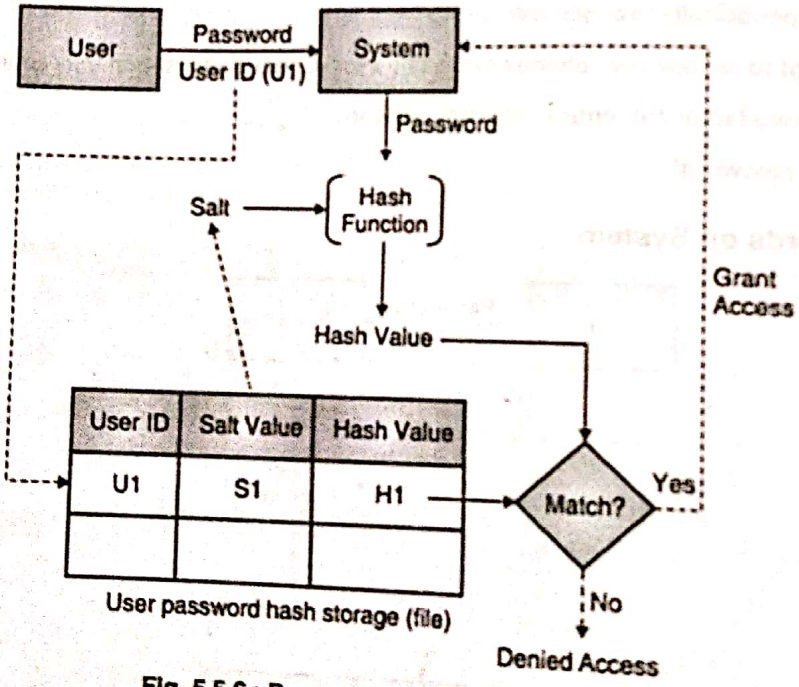


Fig. 5.5.6 : Password verification process

- During the password verification process, the user again provides the user ID and the password she set during the sign-up process.
- Corresponding to the provided user ID, the salt value, used at the time of sign up process, is picked up from the user password hash storage file. The salt value from the file and the user provided password are then fed to the same hashing function. The hashing function computes the hash value with the provided inputs.
- This hash value is then compared with the hash value stored in the user password hash storage file. If the two hash values match, the user provided the right password and is successfully authenticated to the system and is granted access. If the hash values do not match, the authentication process fails, and the user is denied access to the system. Fig. 5.5.6 explains the verification process.

5.5.8 Attacks, Limitations and Challenges on Password Based Authentication

As you know, password is the weakest form of authentication. There have been several mechanisms to attack on it. Let's learn about them.

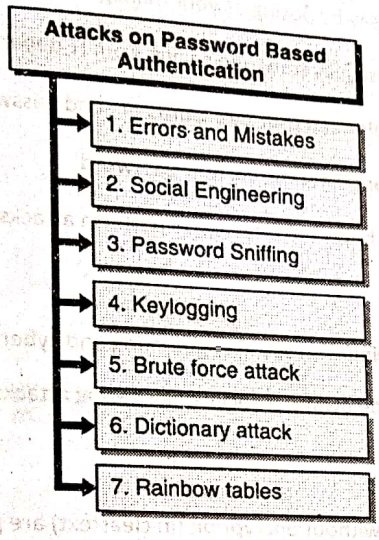


Fig. 5.5.7

1. Errors and Mistakes

- This is probably the most common reason due to which cracking passwords is not that hard. These are common pitfalls against the password creation and general password usage guidelines that you learnt earlier. The common mistakes are
  1. Choosing a weak password that does not align with the password selection guidelines
  2. Writing down passwords
  3. Sharing them with others
  4. Transferring or storing them in cleartext
  5. Using the same password for multiple accounts
  6. Choosing common passwords (something as in the top 25 worst passwords list that you saw earlier)
- Such errors and mistakes make it quite easy to crack password-based authentication and gain unauthorised access to the system.

Protection Mechanism

1. Avoid common pitfalls and adhere to the password selection and password usage guidelines.
2. Use two-factor authentication to ensure that the authentication is not processed only based on the provided password. Another form of authentication would be required before any access is granted to the system.

2. Social Engineering

As you learnt earlier, social engineering attacks can be carried out on the users. Users can be tricked to

1. Reveal the current password (say by posing as your manager needing access to something very quickly)
2. Reset the password (say by posing as someone calling from IT department and asking for quick reset)
3. Ask to install software that could read user passwords or send password hash storage file to the attacker
4. Click on links that could reset or reveal the current password
5. Or carry out anything else that helps the attacker to perform attacks on the password-based authentication

Protection Mechanism

1. Periodic user education and training on social engineering and cybersecurity.
2. Users should be repeatedly made aware of social engineering attacks and how not to get trapped.

3. Password Sniffing

- Passwords that are transferred without encryption (in cleartext) are prone to sniffing.

Definition : *Packet Sniffing is the act of intercepting (capturing) of network traffic and logging it for further analysis.*

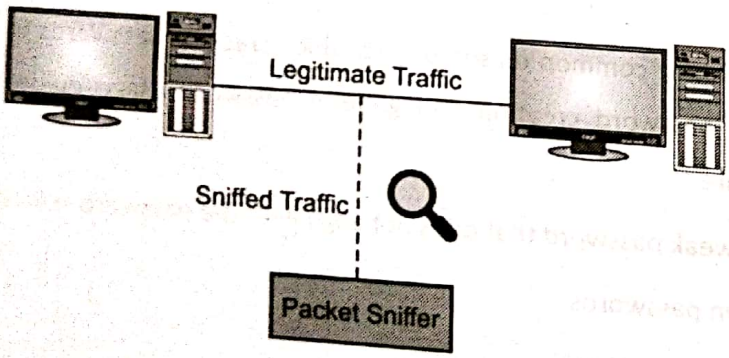
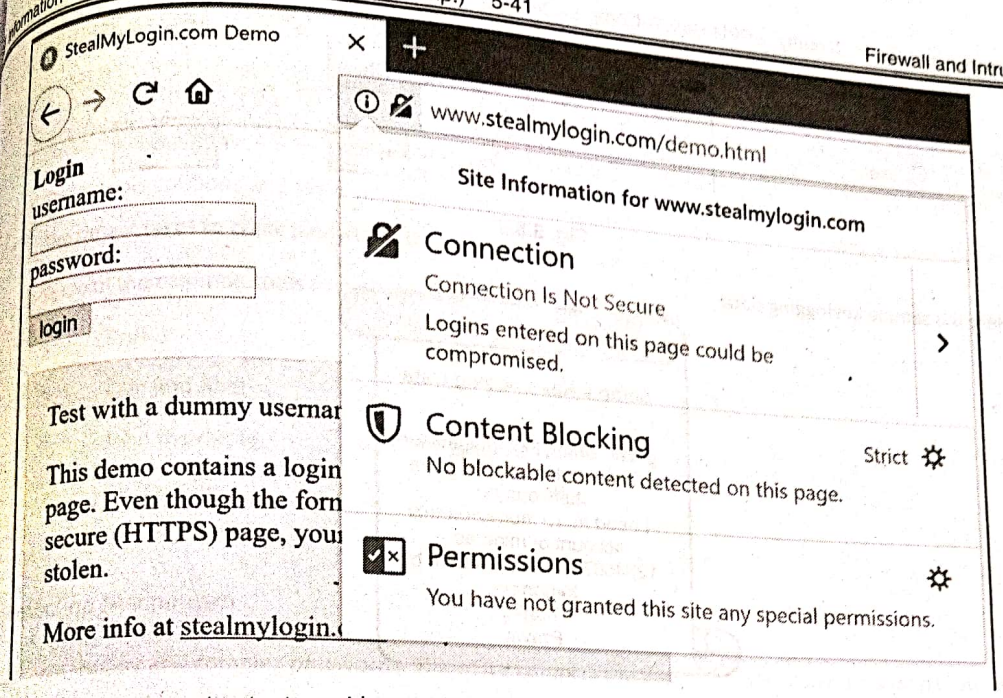


Fig. 5.5.8

- Packet sniffing is carried out using either software programs or hardware devices. These are called packet analysers or just packet sniffers. Wireshark is one of the most widely used Packet Sniffing program.
- Hence, as you see, as a best practice, never send or store your passwords in cleartext. These days browsers give you a warning when you are entering a password in a HTTP (and not HTTPS) site. The HTTP protocol sends information to the webserver in cleartext. You should not ignore such browser warnings. Avoid entering sensitive information such as password on such sites.



- You cannot, to a great extent, avoid network sniffing attempts since it is external to the network and mostly beyond your control. For example, how do you protect someone with Wi-Fi analysing device to capture your Wi-Fi packets? However, some of the common protection mechanisms to reduce the impact from sniffing are as follows.

Protection Mechanism

1. **Encryption :** Use TLS, VPN tunnels or application level encryption wherever possible to avoid sending and receiving data in plaintext.
2. **Use secure protocols :** Discourage the use of protocols such as HTTP, FTP, TELNET, RPC and prefer using secure protocols such as HTTPS, FTPS, SSH, etc.
3. **Isolate and Segment Networks :** Design your network in such a way that the sensitive systems are adequately segmented from rest of the network.

4. Keylogging

- Keylogging is an old technique where the attacker installs a hardware device or a malware on your system that captures all your keystrokes.

Definition : *Keyloggers or simply keystroke loggers are malware programs that capture your keystrokes on the keyboard.*

- The simple idea behind keylogging is that any user would definitely type sensitive information at some point in time. By identifying patterns in the key logged data, the attacker can search for sensitive information. Keyloggers are generally part of other virus or trojan malwares that secretly transmit logged key data to the attacker.

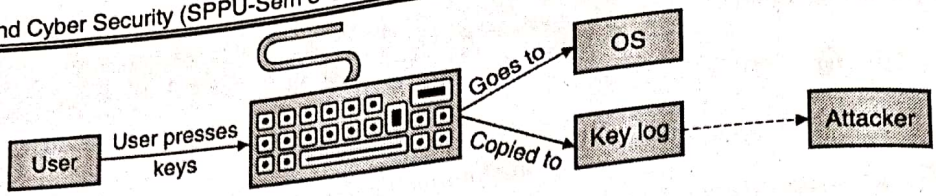
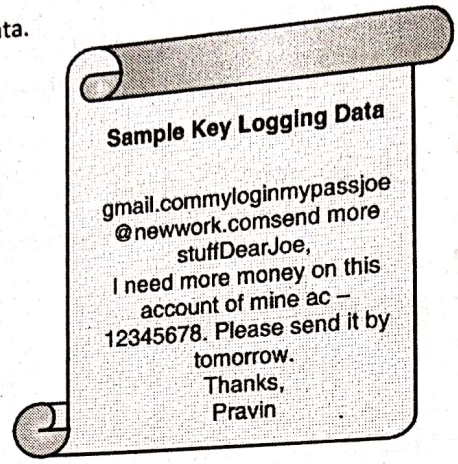


Fig. 5.5.9

- Here is a sample keylogging data.



- Can you identify some patterns and figure out some of my sensitive information? What If I told you that the data belongs to my activity of sending an email via my Gmail account to a friend? Can you figure out the sensitive information?
- Let me help you. Assume the following sequence of activities.
  1. I opened my favourite browser Mozilla and typed "gmail.com".
  2. The Gmail login page appeared and then I typed my username "mylogin" and my password "mypass".
  3. Then, I clicked on *compose* and typed the email address of my friend "joe@newwork.com".
  4. Then, I typed my email subject "send more stuff".
  5. Then, I composed my email and hit the *send* button.
- Does it sound simple but scary? That's precisely what keylogging is. All keys are captured as you press them. Some modern and advanced keyloggers also capture mouse clicks (x and y coordinates of the screen or pixel position information on the screen) along with screenshots.

Protection Mechanism

1. Regularly use anti-malware programs to scan your systems.
2. Set multi-factor authentication to avoid account hijack even if your password is compromised.
3. Use virtual keyboard specially when dealing with financial related logins.
4. Regularly check your installed program inventory to identify any potential malware.
5. Regularly check the services running on your system to identify any potential malware service.
6. Physically inspect your system wires and attachments to ensure that there are no unwanted hardware plugins.

5. Brute Force Attack

***Definition** : Bruce force attack is performed using tools that cycle through various combinations of characters until a successful combination of characters is found.*

- These tools continuously send various character combinations in the hope of finding a successful combination. The time it takes to crack passwords this way is usually large.
- Some of the common tools to carry out the brute force attack are

1. Brutus
2. RainbowCrack
3. Cain and Abel
4. Wfuzz
5. John the ripper
6. Medusa
7. Aircrack-ng

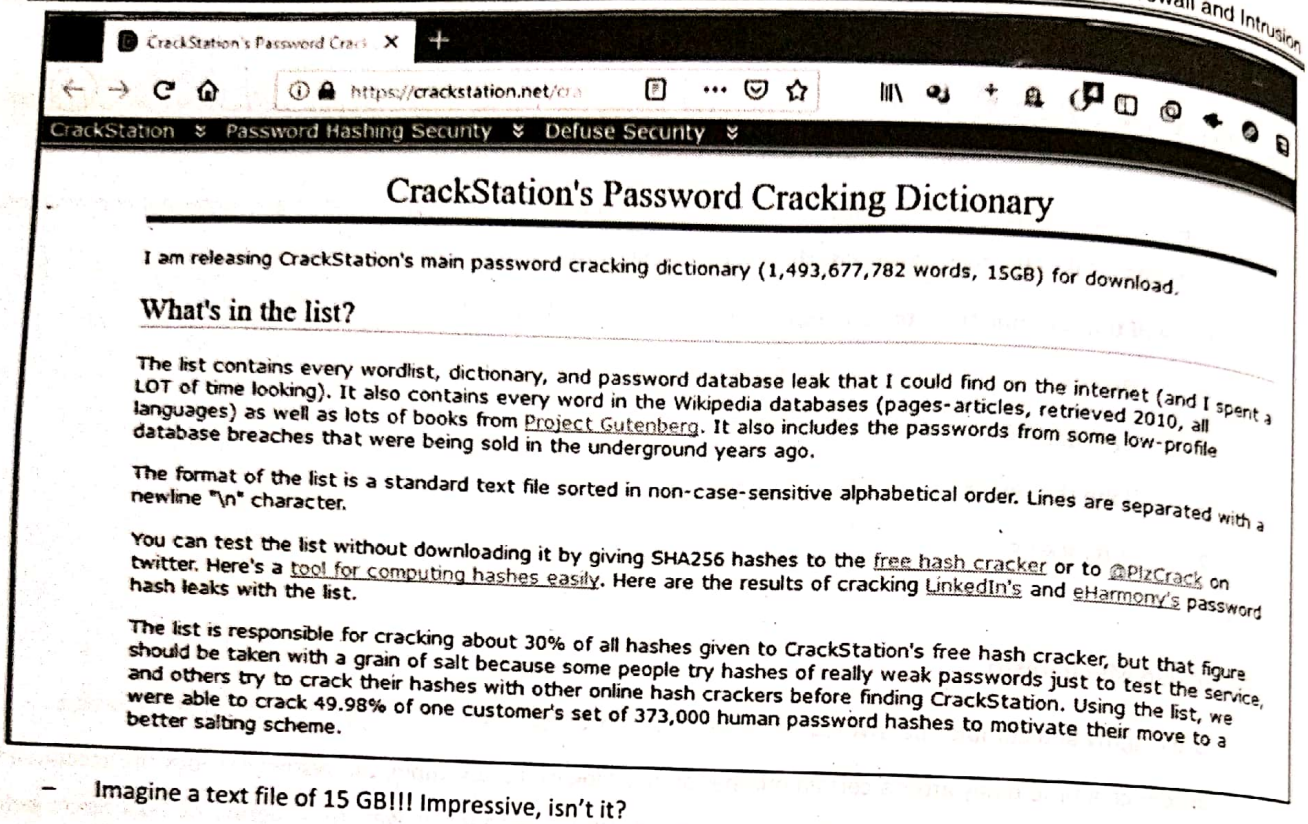
Protection Mechanism

1. Use lengthy and complex passwords. The more complex the password, the longer it takes to brute force it.
2. Introduce a time delay after a certain number of failed logins. For example, the system can lock the account after 3 failed attempts and requires administrator to unlock the account or wait for a period of 1-24 hours before automatically unlocking the account.
3. Send login notification. Whenever your account is attempted to log in, you could be notified via SMS or email. You could then take corrective actions if the login activity is not found to be legitimate.
4. Protect the password hash storage files from any unauthorised access or download. The attacks can be carried out on the files thus bypassing all forms of controls.

6. Dictionary Attack

***Definition** : Dictionary attack is a variation of brute force attack where instead of trying random character combinations, dictionary words and their variations are tried out for cracking the password.*

- Dictionaries could be used from the user's natively spoken language or the multiple languages that a user may know. Apart from the regular dictionaries, there are other dictionaries which are collections from the hacker community.
- When a large user account database is stolen, such information is available with the hacker community that uses the information to form a listing of new passwords that people might be using.
- To give you a feel of how large a dictionary could be, take a look at the following snapshot.



- Imagine a text file of 15 GB!!! Impressive, isn't it?
- Dictionary attack might take lesser time than the actual brute force attack since not all combinations are tried. Also, if the user has not chosen a dictionary-based password, then the dictionary attack may not be able to crack it.

Protection Mechanism

1. Follow password selection guidelines and choose a complex and non-dictionary password.
2. Introduce a time delay after a certain number of failed logins. For example, the system can lock the account after 3 failed attempts and requires administrator to unlock the account or wait for a period of 1-24 hours before automatically unlocking the account.
3. Send login notification. Whenever your account is attempted to log in, you could be notified via SMS or email. You could then take corrective actions if the login activity is not found to be legitimate.
4. Protect the password hash storage files from any unauthorised access or download. The attacks can be carried out on the files thus bypassing all forms of controls.

7. Rainbow Tables

Definition : Rainbow tables consist of passwords in the hash format.

- Rainbow tables are minor variations of dictionaries used in the dictionary attacks. Unlike dictionary attacks that use the pre-defined character combinations, rainbow tables contain pre-computed hash values for every possible character combination in those dictionaries.

That way, you need not waste system resources and time in computing the hash values (again and again) and then comparing with the hash value on the system (or the stolen user password hash storage file).

Table 5.5.2 is a simplistic rainbow table.

Table 5.5.2 : A simple rainbow table

Character Combination	SHA-1 Hash
NewPassword	735b535bd8148743e2e19b08c7db8ea60142a0b3
MyPassword	daa1f31819ed4928fd00e986e6bda6dab6b177dc
Iamhappy	b0177566098815e0555830d7a68d2352614fe9bc
Apple@123	d88f1b3fb247e4ad3f6b821e4da902d1c91f0864

- Your problem is simplified. All you need to find then is any hash value in the rainbow table that could match with the system's hash value.
- When such a match is found, you can find the corresponding character combination from the rainbow table. This character combination is the cracked password.

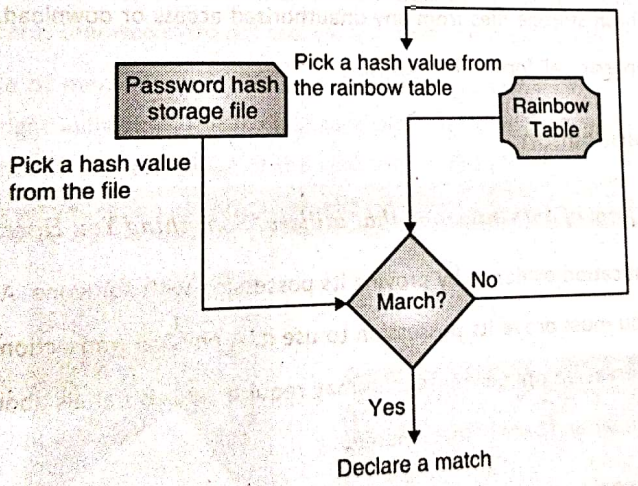
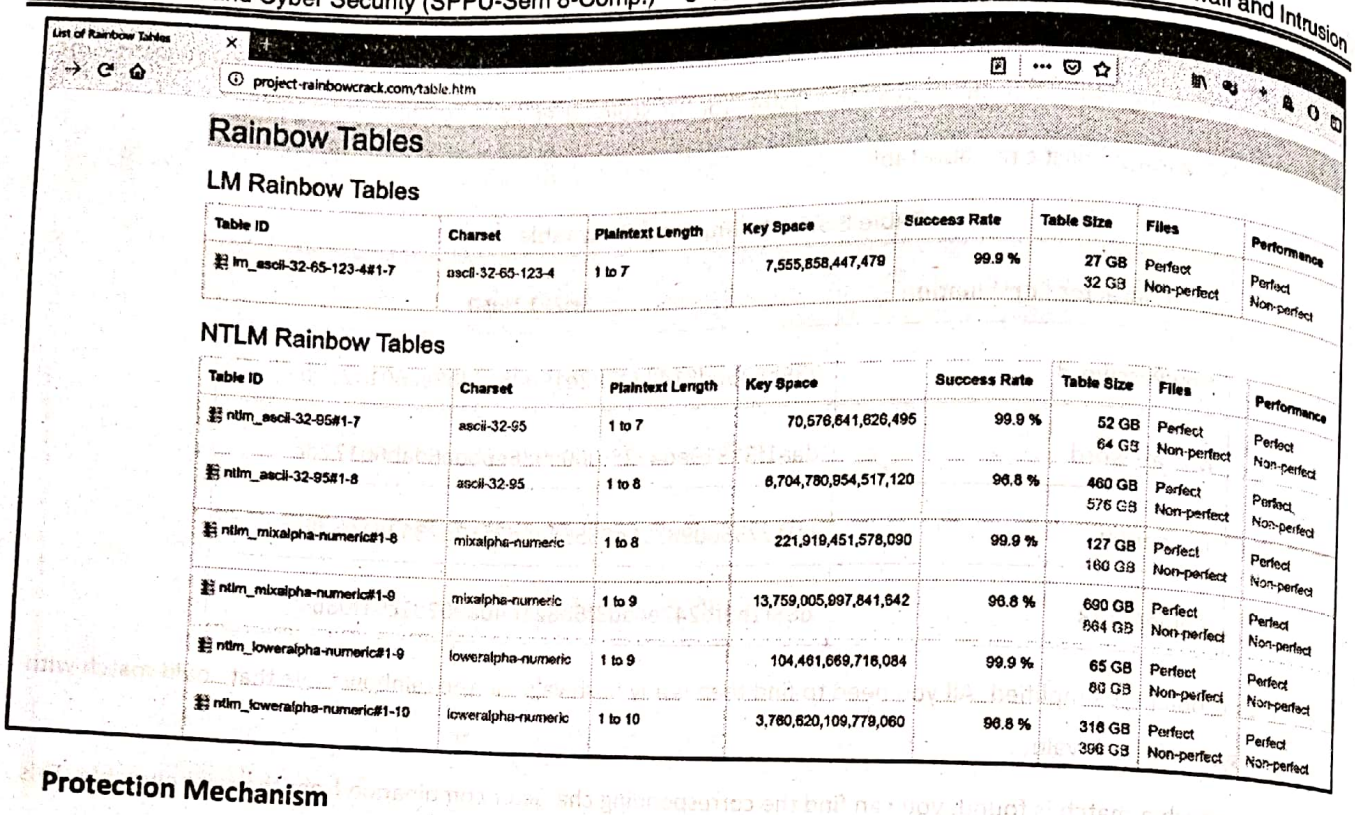


Fig. 5.5.10

- Rainbow tables are much faster when compared with the brute force or dictionary attacks. RainbowCrack is one of the tools that use rainbow tables to crack passwords.
- The rainbow tables are also huge in size as they contain several hash values for dictionary words.



List of Rainbow Tables

project-rainbowcrack.com/table.htm

Rainbow Tables

LM Rainbow Tables

Table ID	Charset	Plaintext Length	Key Space	Success Rate	Table Size	Files	Performance
lm_ascii-32-65-123-4#1-7	ascii-32-65-123-4	1 to 7	7,555,858,447,479	99.9 %	27 GB 32 GB	Perfect Non-perfect	Perfect Non-perfect

NTLM Rainbow Tables

Table ID	Charset	Plaintext Length	Key Space	Success Rate	Table Size	Files	Performance
ntlm_ascii-32-95#1-7	ascii-32-95	1 to 7	70,576,641,626,495	99.9 %	52 GB 64 GB	Perfect Non-perfect	Perfect Non-perfect
ntlm_ascii-32-95#1-8	ascii-32-95	1 to 8	6,704,780,954,517,120	96.8 %	460 GB 576 GB	Perfect Non-perfect	Perfect Non-perfect
ntlm_mixalpha-numeric#1-8	mixalpha-numeric	1 to 8	221,919,451,578,090	99.9 %	127 GB 160 GB	Perfect Non-perfect	Perfect Non-perfect
ntlm_mixalpha-numeric#1-9	mixalpha-numeric	1 to 9	13,759,005,997,841,642	96.8 %	690 GB 864 GB	Perfect Non-perfect	Perfect Non-perfect
ntlm_loweralpha-numeric#1-9	loweralpha-numeric	1 to 9	104,461,669,716,084	99.9 %	65 GB 80 GB	Perfect Non-perfect	Perfect Non-perfect
ntlm_loweralpha-numeric#1-10	loweralpha-numeric	1 to 10	3,760,620,109,779,060	96.8 %	316 GB 396 GB	Perfect Non-perfect	Perfect Non-perfect

Protection Mechanism

1. Follow password selection guidelines and choose a complex and non-dictionary password.
2. Protect the password hash storage files from any unauthorised access or download. The attacks can be carried out on the files thus bypassing all forms of controls.

5.5.9 Token Based Authentication

Definition : *Token is a form of authentication that utilizes "Something You Have" type of authentication.*

- A token serves as an authentication evidence by proving its possession with someone. A simple example of token is your Debit or Credit Card. You must prove its possession to use it. A physical transaction may require a swipe of the card to prove possession whereas an online transaction may require various details about the card such as the card number, expiry date, CVV number and card holder's name.
- There are various types of token devices used in different ways to provide authentication.

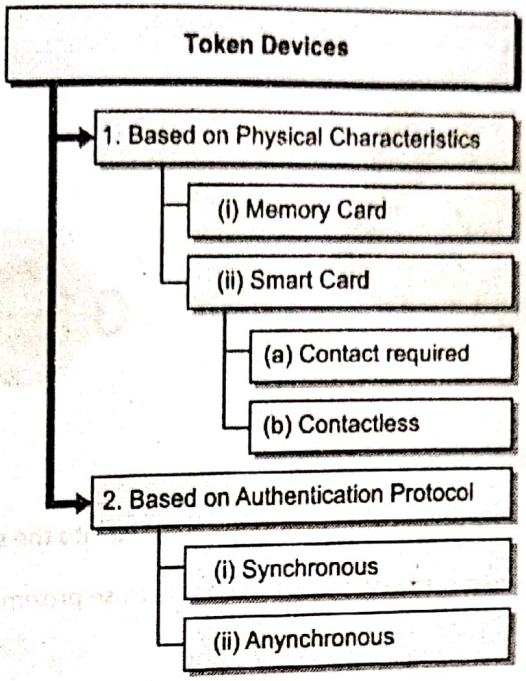


Fig. 5.5.11

1. Memory Card

Definition : Memory cards physically contain the required authentication information.

- It usually has a magnetic stripe (or a chip) and the authentication information is contained in this stripe (or chip). You require a card reader using which you can swipe the card (very much like your Debit Card) and authenticate yourself. Memory cards themselves do not possess any form of processing power.
- Sometimes, the use of memory cards also requires a password or a PIN to validate that it is actually in the possession of the right individual. This is done to avoid circumstances where stolen cards can be used just by their possession without the knowledge of the card owner. The use of a PIN or a password along with the card possession, makes authentication as two-factor authentication.

2. Smart Card

Definition : Smart card has processing capacity that can be used to provide authentication information.

- Unlike memory cards, smart cards have processing capacity and may also have display and buttons. It may or may not require a smart card reader.

- There are two types of smart cards.
  1. **Contact required :** The smart card needs to be inserted into the smart card reader for authentication.
  2. **Contactless :** The smart card just needs to be in the close proximity of the reading device.

C. Synchronous Tokens

Definition : *A synchronous token device synchronizes (works tightly coupled) with the authentication service for providing authentication information.*

- The synchronisation could be based on time or a counter. RSA SecurID is an example of time-based synchronous token.

- If the synchronization is time-based, the token device and the authentication service must hold the same time within their internal clocks.
- The time value on the token device and a secret key (pre-loaded in the token device and the authentication server also has a copy of it) are used to create the one-time password, which is displayed to the user.
- The user enters this value and a user ID into the computer, which then passes them to the server running the authentication service.
- The authentication service compares it to the value it expected. If the two match, the user is authenticated and allowed to use the computer and resources.
- If the token device and authentication service use counter-based synchronization, the user will need to initiate the creation of the one-time password by pushing a button on the token device.
- This causes the token device and the authentication service to advance to the next authentication value. The user enters this resulting value along with a user ID to be authenticated.

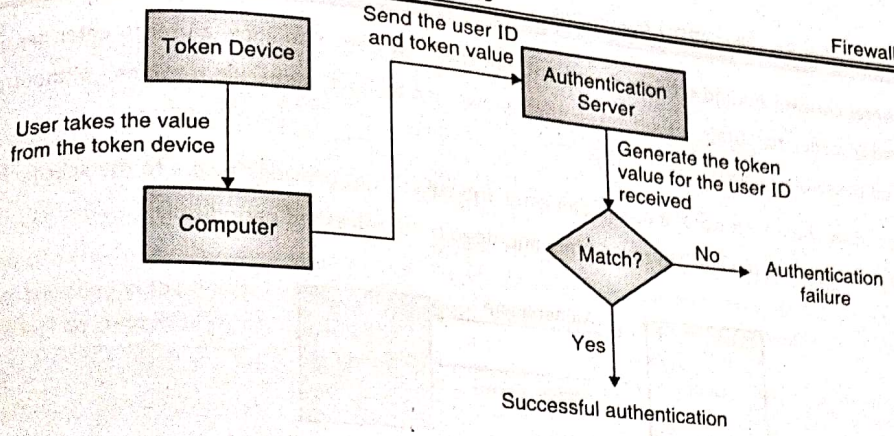


Fig. 5.5.12

D. Asynchronous Tokens

Definition : *Asynchronous token device uses challenge-response mechanism for providing authentication information.*

- Asynchronous token is not time dependent. It uses a challenge-response mechanism (very much like handshake protocols) for providing authentication information.
- The authentication server sends the user a challenge which is usually a random value. The user needs to enter this value into the token device. The token device encrypts the value entered and returns a value that the user uses as a one-time password. The user sends this value along with the user ID to the authentication server. If the authentication server can decrypt the value and gets the same challenge value that it sent earlier, the user is authenticated.

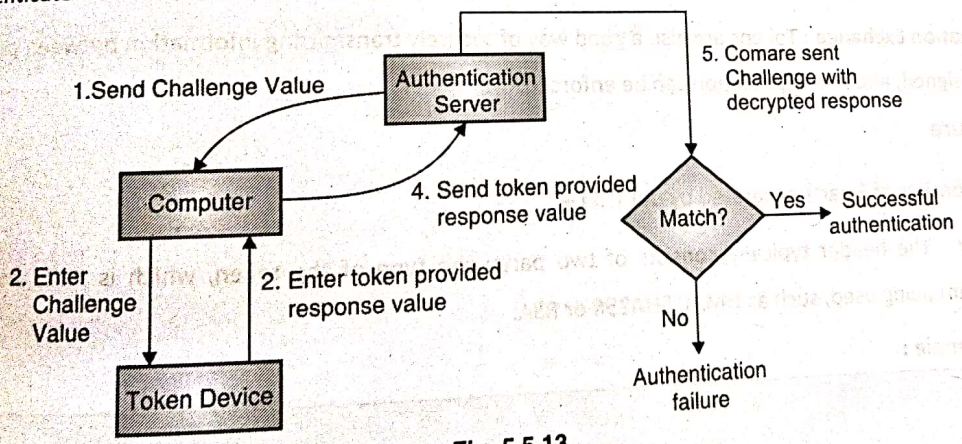


Fig. 5.5.13

E. Software Tokens

- Similar to physical tokens there also exist software-based tokens. Software tokens are used for security claims (authorisation) between two communicating parties. One such implementation, based on RFC 7519, is JSON Web Token or JWT (pronounced Jot).

Definition : *JSON Web Token (JWT) is a compact, URL-safe means of representing claims to be transferred between two parties.*

- The general concept behind a token-based authentication system is to allow a user to enter her user name and password in order to obtain a web token that allows her to fetch a specific resource - without using her user name and password again.
- Once her token has been obtained, she can offer the token that provides access to the specific resources for a time period. Tokens come with an expiry time and need to be refreshed periodically.

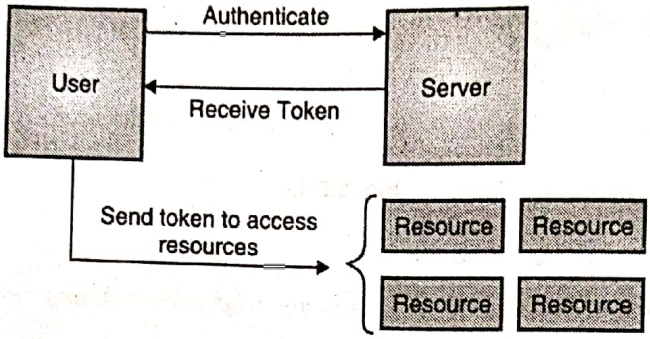


Fig. 5.5.14

Usage of software tokens

Two common scenarios where software tokens are used are

1. **Authorisation :** Once the user is logged in, each subsequent request will include the token, allowing the user to access resources that are permitted with that token. The user does not require re-authentication several times to continue accessing the permitted resources.
2. **Information Exchange :** Tokens are also a good way of securely transmitting information between parties. Tokens can be signed, and non-repudiation can be enforced.

F. JWT Structure

- A JWT consists of 3 parts separated by dot (".") –

1. **Header :** The header typically consists of two parts: the type of the token, which is JWT, and the signing algorithm being used, such as HMAC SHA256 or RSA.

For example :

```
{
 "alg": "HS256",
 "typ": "JWT"
}
```

- Here algorithm is HMAC SHA 256 written HS256 in short. Token type is JWT.

2. **Payload :** The second part of the token is the payload, which contains the claims. Claims are statements about an entity (typically, the user) and additional data.

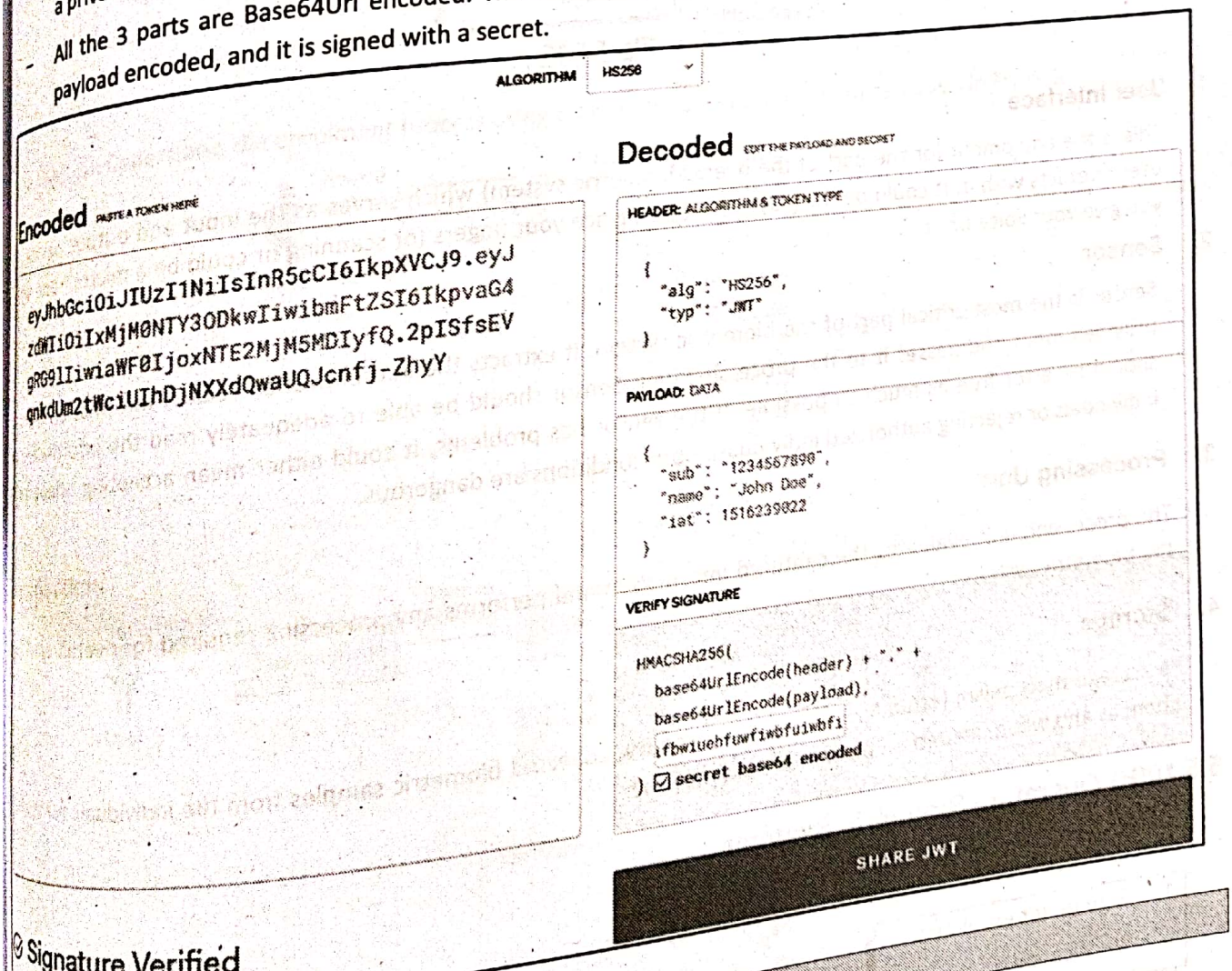
For example :

```
{
"sub": "1234567890",
"name": "John Doe",
"admin": true
}
```

3. **Signature :** To create the signature part you have to take the encoded header, the encoded payload, a secret, the algorithm specified in the header, and sign that. For example if you want to use the HMAC SHA256 algorithm, the signature will be created in the following way:

```
HMACSHA256(
base64UrlEncode(header) + "." +
base64UrlEncode(payload),
secret)
```

- The signature is used to verify the message wasn't changed along the way, and, in the case of tokens signed with a private key, it can also verify that the sender of the JWT is who it says it is.
- All the 3 parts are Base64Url encoded. The following example shows a JWT that has the previous header and payload encoded, and it is signed with a secret.



Note : You can go to https://jwt.io/ and try out JWT.

5.5.10 Biometric Based Authentication

- The biometric-based authentication relies upon "Someone You Are" and "Something You Do". "Something You Are" utilizes physical characteristics of your body whereas "Something You Do" utilizes your behavioural characteristics.
- Biometric is the most expensive way of authentication. It requires specialized readers, processing, storage and comes with errors.

5.5.10(A) Components of Biometric Systems

Any typical biometric system comprises of the following components.

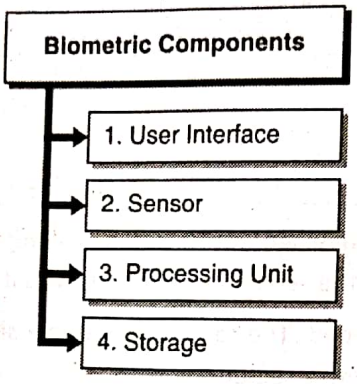


Fig. 5.5.15

1. User Interface

This is the equipment (or the part of the overall biometric system) which serves as the input and output device. The user interacts with it. It could be the glass where you place your fingers for scanning or could be a microphone where you give your voice print.

2. Sensor

Sensor is the most critical part of the biometric system. It extracts the authentication related information from the provided input and passes it to the processing unit. Sensor should be able to adequately read the information and should be error free as much as possible. If the sensor has problems, it could either mean accepting unauthorized individuals or rejecting authorized individuals. Both conditions are dangerous.

3. Processing Unit

The processing unit evaluates the captured information and performs any processing required for overall working of the biometric system.

4. Storage

A storage mechanism (or unit) is required to keep the collected biometric samples from the individuals for matching them as and when needed for authentication purpose.

5.5.10(B) Operating Biometric Systems

Like passwords, biometric systems also have two phases.

1. Sign up (or enrolment)
2. Verification

Enrolment

The enrolment process involves collecting the biometric sample from the individual. Remember your Aadhar card enrolment? Your physical presence is required. You provide the biometric sample (one or multiple times) based on the type of the biometric device.

If it is a fingerprint scanner, you provide your fingerprints. If it is a retina scanner, you look through an eye scanner. Once your sample is collected, the information that can be used for authentication is extracted from it. The information is digitized in the binary format and is stored for future use.

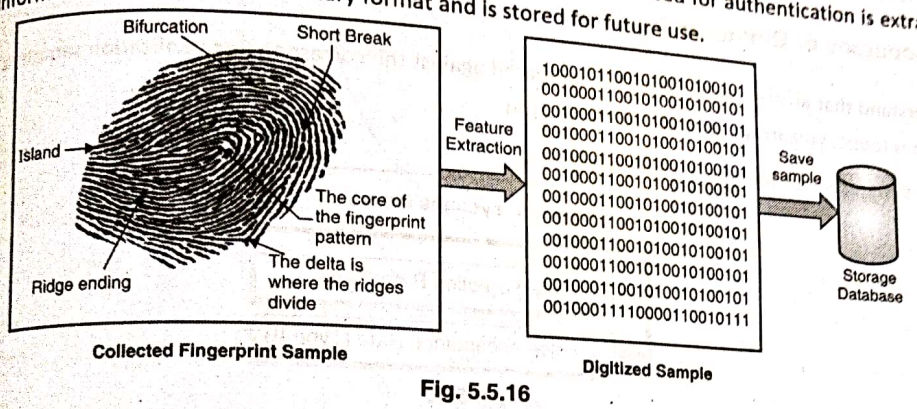


Fig. 5.5.16

You can understand the enrolment process using a simplistic schematic diagram as shown in Fig. 5.5.17.

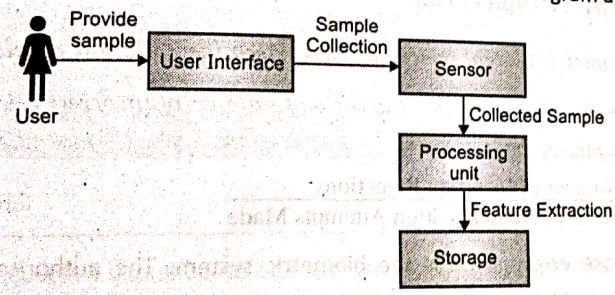


Fig. 5.5.17

Verification

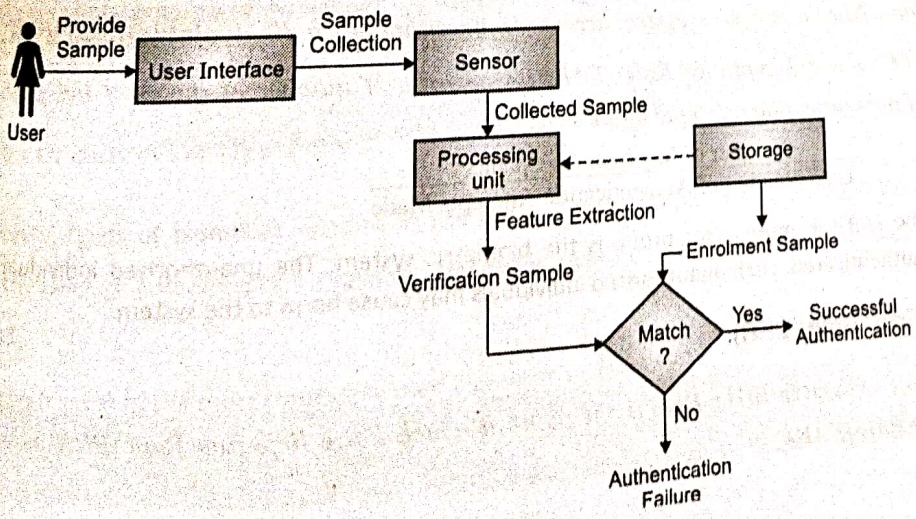


Fig. 5.5.18

(Ref No. - 3673/2019-CO/L & 8811/2019-CO/L)

- The verification process is quite straightforward. Your physical presence is again required. You offer the same type of sample that you provided during the enrolment process. For example, if you enrolled with fingerprints, you would need to provide fingerprints again for verifying.
- Your provided verification sample is compared against the enrolment sample based on your identity. Once the two samples match, you are successfully verified. If the samples do not match, the authentication is rejected. This might be due to some error. In that case, you can retry providing the sample again.

5.5.10(C) Accuracy of Biometric Systems

- You understand that an enrolment sample is compared against the corresponding verification sample to find a match. If a match is found, you are successfully authenticated.
- Biometric systems are prone to two types of errors :

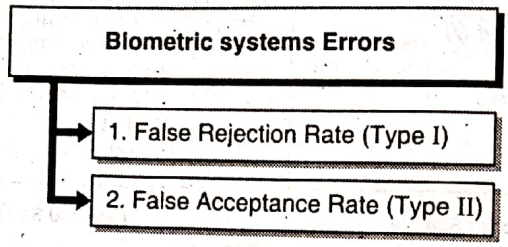


Fig. 5.5.19

1. False Rejection Rate (FRR) or Type I error

✍ ***Definition :*** *When a biometric system rejects an authorised individual it is called Type I error.*

✍ ***Definition :*** *False Rejection Rate (FRR) is the ratio of number of incorrect rejections to the total number of authentication attempts made.*

$$FRR = \frac{\text{Number of Incorrect Rejections}}{\text{Total Number of Authentication Attempts Made}}$$

The higher the FRR the more error prone is the biometric system. The authorised individuals would likely get frustrated from repeated attempts required to pass authentication.

2. False Acceptance Rate (FAR) or Type II Error

✍ ***Definition :*** *When a biometric system accepts an unauthorised individual it is called Type II error.*

✍ ***Definition :*** *False Acceptance Rate (FAR) is the ratio of number of incorrect acceptances to the total number of authentication attempts made.*

$$FAR = \frac{\text{Number of Incorrect Acceptances}}{\text{Total Number of Authentication Attempts Made}}$$

The higher the FAR the more error prone is the biometric system. The unauthorised individuals would likely get successfully authenticated. Such unauthorised individuals may cause harm to the system.

3. Crossover Error Rate (CER)

✍ ***Definition :*** *Crossover Error Rate is the point at which False Rejection Rate (FRR) is equal to the False Acceptance Rate (FAR).*

This means that the biometric system is not more likely to produce one type of error than the other type. If the system rejects too many authorised individuals the FRR would be high. Similarly, if the system accepts too many unauthorised individuals the FAR would be high. CER is also called as Equal Error Rate (EER).

When comparing various biometric systems, choose the one with lower CER value. The lower the CER value the more accurate the biometric system is. For example, a biometric system having CER value 3 would be more accurate than the biometric system having CER value of 5.

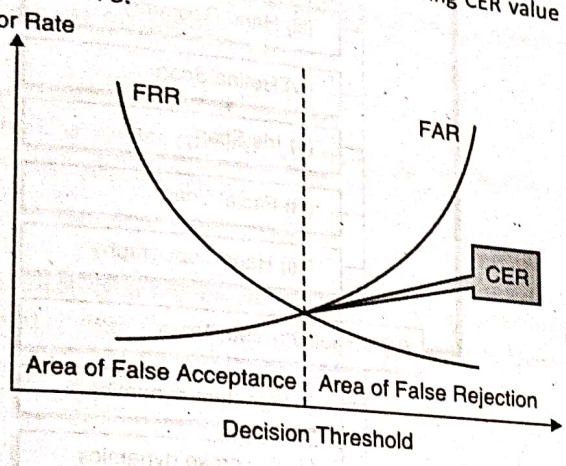


Fig. 5.5.20

5.5.1: A biometric system rejects 2 authorised attempts out of 10. Calculate its FRR.

Soln:

$$FRR = \frac{\text{Number of Incorrect Rejections}}{\text{Total Number of Authentication Attempts Made}}$$

Hence, $FRR = 2/10 = 0.2$

5.5.2: A biometric system accepts 4 unauthorised attempts out of 10. Calculate its FAR.

Soln:

$$FAR = \frac{\text{Number of Incorrect Acceptances}}{\text{Total Number of Authentication Attempts Made}}$$

Hence, $FAR = 4/10 = 0.4$

5.10(D) Types of Biometric Systems

There are several types of biometric systems in use today. These systems are used depending upon the type of application, the level of authentication required, processing speed and authentication accuracy (Type I errors and Type II errors).

Let us learn about some of the most commonly used biometric systems.

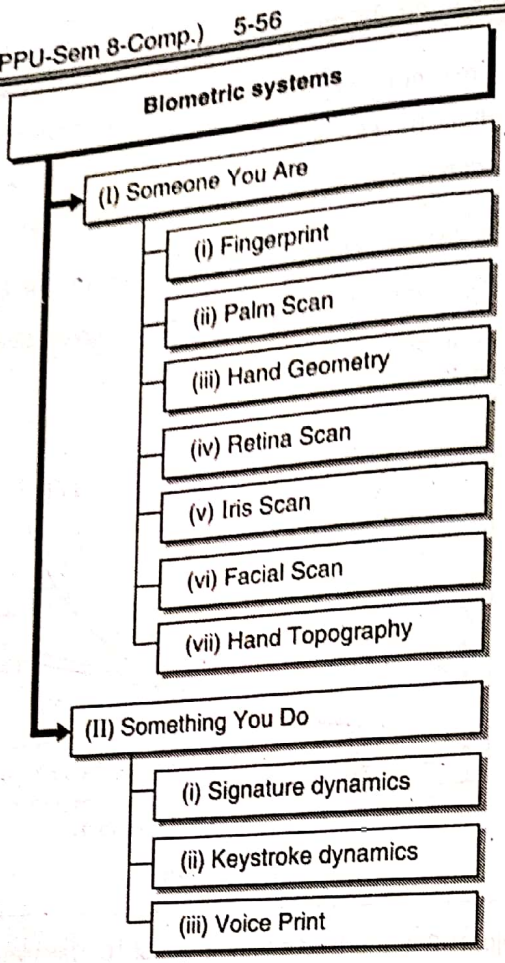


Fig. 5.5.21

1. Fingerprint

Fingerprint is one of the most commonly used type of biometric system. Your fingerprint is made up of several curves and endpoints. These when combined together are unique enough to identify you.

2. Palm Scan

You would have seen this in several movies. Your palm contains several curves, lines, endpoints, folds and textures that can uniquely identify you. Unlike fingerprints, you need to place your entire palm on the biometric sensor to enroll and verify yourself.

3. Hand Geometry

Your hand holds several key attributes such as shape, length, width, size etc. These attributes could fulfil biometric requirements to provide authentication information.

4. Retina Scan

Retina scan involves reading the blood-vessel pattern of retina on the backside of the eyeball. This pattern is unique enough to identify you. You are required to look into an eye scanner (fitted with a high-resolution camera). The scanner captures the pattern and stores it for authentication.

5. Iris Scan

The iris is the coloured portion of the eye that surrounds the pupil. The iris has unique patterns, rifts, colours, rings, coronas, and furrows. When you look into an eye scanner, these characteristics are captured and stored for authentication. Iris scan is the most accurate form of biometric system.

Facial Scan

Your face can be digitally identified as well. It has got several key features such as nose ridges, eye widths, chin shape, forehead size, bone structures, etc. These features are extracted during a facial scan and stored for authentication.

Hand Topography

Hand topography captures overall hand shape and structure. The peaks and valleys in your hand provide biometric traits that can be used for authentication. The entire hand is placed on a scanner and there are several cameras that capture the hand information from various angles.

Signature Dynamics

You sign at a particular speed making similar strokes each time. The signing process generates electrical signals that can be captured to provide biometric authentication.

Keystroke Dynamics

Keystroke dynamics capture electrical signals when you type using a keyboard. You type at a particular speed applying specific key pressure, timing and rhythm for each key that you press. These dynamics can be captured and used for authentication.

Voice Print

Aren't you using Alexa, Siri or Ok Google yet? How does that work? Basically, you train the device to recognize your voice pattern. Your voice pattern has specific pitch, tone, amplitude and frequency that can be used to create a voice print (similar to fingerprint). These attributes are stored and when you are required to authenticate you are asked to speak a set of words or sentences to capture your voice print and compare it with the previously stored voice print information.

Comparison of Biometric Systems

Sr. No.	Biometric System	Processing Speed	Accuracy	Ease of Enrolment
1.	Fingerprint	High	High	High
2.	Palm Scan	Medium	High	Medium
3.	Hand Geometry	Low	Medium	Low
4.	Retina Scan	Medium	High	High
5.	Iris Scan	Medium	High	High
6.	Facial Scan	Medium	Low	Medium
7.	Hand Topography	Low	Low	Low
8.	Signature dynamics	Medium	Medium	Medium
9.	Keystroke dynamics	Medium	Medium	Medium
10.	Voice Print	Medium	Medium	Medium